Limit Theorems for Two Classes of Markov Processes

Timofei Prasolov

Submitted for the degree of Doctor of Philosophy

Heriot-Watt University
Department of Actuarial Mathematics and Statistics
School of Mathematical and Computer Sciences

July 2019

Abstract

The thesis includes two Parts that analyse the asymptotic behaviour of two multicomponent stochastic processes. In both cases, the components of the processes are highly dependent, however dynamics of two processes are significantly different.

Part I is devoted to the study of hierarchical models with local dependence: behaviour of the $(i+1)$'st component is influenced by the i'th component only. These processes are either null-recurrent or transient and, therefore, do not possess limiting distributions. We analyse the structure of these processes and obtain limit theorems under normalisation.

Part II deals with another type of models that arise in the neural systems. We consider symmetric models: any permutation of the coordinates has the same type of dynamics. We analyse the structure of these processes and conditions for positive recurrence.

Acknowledgements

I would like to sincerely thank my supervisors for their unbounded support and understanding during last four years. I had some troubles with living abroad and I was struggling with a number of health issues. Without their advices and support I would not go through it.

I am deeply grateful to my first supervisor, Sergey Foss. His deep mathematical knowledge, infinite amount of fun stories and overwhelming care for people is beyond praise. I want to thank my second supervisor, Seva Shneer. His guidance and understanding helped me with a number of difficult situations.

I want to thank Dima Korshunov, Denis Denisov and Vitali Wachtel for their help during the research. Special thanks to Thibaud Taillefumier for the introduction to neuron networks.

Contents

Part I. Limit theorems for a class of hierarchical models **1**

Section 1. Introduction and main results **1**

§ 1.1. Markov chain and Markov process 1

1.1.1. Harris positive recurrence 2

§ 1.2. Cat-and-Mouse Markov chain and its appearances in practice . . . 3

§ 1.3. Heavy-tailed distributions and related results 6

1.3.1. Subexponential distributions 6

1.3.2. Tail asymptotics for randomly stopped sums 8

1.3.3. Hitting times for random walks and renewal processes 9

§ 1.4. Lévy processes, Brownian motion and local time 11

§ 1.5. Weak convergence for processes taking values in $D[[0,\infty),\mathbb{R}]$ . . . 11

§ 1.6. Scaling results for the mouse on $\mathbb{Z},\mathbb{Z}^2$ or $\mathbb{Z}^+$ 13

§ 1.7. Stable laws 15

§ 1.8. General overview of our results 16

§ 1.9. Related results 18

§ 1.10. Structure of Part I 21

§ 1.11. "Standard" Cat-and-Mouse Markov chain on $\mathbb{Z}$ $(C \to M)$ 22

§ 1.12. Cat-and-Mouse model with a general jump distribution of the mouse $(C \to M)$ 22

§ 1.13. Dog-and-Cat-and-Mouse model $(D \to C \to M)$ 24

§ 1.14. Linear hierarchical chains $(X_1 \to X_2 \to \ldots \to X_N)$ of length N . 25

Section 2. Linear hierarchical chain of length N **27**

Section 3. The Cat-and-Mouse model **31**

Section 4. The Dog-and-Cat-and-Mouse model **37**

§ 4.1. Comments on the CM model 37

§ 4.2. Auxiliary continuous-time random walk in the DCM model 38

§ 4.3. Distribution of random variable $J_1^{(3)}$ 39

§ 4.4. Distribution of random variable $Y_1^{(3)}$ 41

§ 4.5. Proof of Theorem **1.21** 42

Part II. Stability of neural networks **44**

S e c t i o n 5. Introduction and main results **44**

§ 5.1. Modelling of neural networks 44

5.1.1. Perfect integrate-and-fire neuron model 45

§ 5.2. General overview of our results 47

5.2.1. Fluid model and positive recurrence 48

5.2.2. Ergodicity of processes admitting embedded Markov chains . 49

5.2.3. Related results 52

§ 5.3. Structure of Part II 54

§ 5.4. Model and results 54

S e c t i o n 6. Stability of the system with Lévy input **59**

§ 6.1. Fluid model and corresponding auxiliary results 59

§ 6.2. Proofs of auxiliary results 61

§ 6.3. Proof of positive recurrence 65

§ 6.4. Comments on average drifts μ_i and signals b_{ij} 67

§ 6.5. Proof of existence of minorant measure 68

S e c t i o n 7. Spike-analysis in two particular cases **71**

§ 7.1. Constant input 71

§ 7.2. Poisson input 75

S e c t i o n 8. Future research **79**

S e c t i o n A. Appendix I **84**

§ A.1. Proof of Proposition **1.12** 84

§ A.2. Limit theorems for CTRW's 85

S e c t i o n B. Appendix I **87**

§ B.1. Proof of Remark **5.7** 87

§ B.2. Proof of Lemma **6.4** 87

References **89**

Part I. Limit theorems for a class of hierarchical models

SECTION 1
Introduction and main results

§ 1.1. Markov chain and Markov process

In this Subsection we define our terminology regarding Markov chains which we use in this Section.

Definition 1.1. A discrete time stochastic process $\{X_n\}_{n\geq 0}$ taking values in a countable state space $\mathcal{S}$ is a *Markov chain* if and only if

$$\mathbb{P}\{X_{n+1} = j |\ X_n = i,\ X_{n-1} = i_{n-1}, \ldots, X_0 = i_0\} = \mathbb{P}\{X_{n+1} = j |\ X_n = i\}$$

for all $n \geq 0$ and for all $j, i, i_{n-1}, \ldots, i_0 \in \mathcal{S}$.

The Markov chain $\{X_n\}_{n\geq 0}$ is *time homogeneous*, if, for all $i, j \in \mathcal{S}$, there exists p_{ij} such that

$$\mathbb{P}\{X_{n+1} = j |\ X_n = i\} = p(i, j)$$

independently of n. The matrix $P = (p(i, j))_{i,j\in\mathcal{S}}$ is called transition matrix.

We consider only *irreducible* Markov chains, i.e. for every $i, j \in \mathcal{S}$ there exists $n \geq 1$ such that $p^{(n)}(i, j) = \mathbb{P}\{X_n = j |\ X_0 = i\} > 0$. In other words, between any states there is a path which the Markov chain takes with positive chance. We are also interested whether the Markov chain is *aperiodic*. There are different ways to define aperiodicity (see, e.g., Spitzer (1964)). For irreducible Markov chains we can define it as follows: for every state i the greatest common divider of $\{n \geq 1 : p^{(n)}(i, i) > 0\}$ equals one. This property leads to a 'mixing of the system': after enough time if you visit a state you can not concur at which state you started and how long ago.

Definition 1.2. A (probability) distribution $\vec{\pi}$ on $\mathcal{S}$ is **stationary** for the

Markov chain if and only if

$$\vec{\pi}P = \vec{\pi} \qquad \text{or equivalently} \qquad \sum_{i \in \mathcal{S}} \pi_i p(i,j) = \pi_j, \quad j \in \mathcal{S}. \tag{1}$$

Sometimes there is no solution to system (1) which has at least one positive element π_i and has a finite weight $\sum_{i \in \mathcal{S}} \pi_i$. In this case we may talk about *invariant measures* $\vec{\mu}$ on $\mathcal{S}$. For example, in a case of a simple random walk on $\mathbb{Z}$ the transition matrix takes form $p(i, i+1) = p(i, i-1) = 1/2$. Then if we restrict $\vec{\pi}$ to have non-negative elements and finite weight then it is a zero-vector. Nevertheless, a vector $\vec{\mu}$, such that $\mu_i = k > 0$, $i \in \mathbb{Z}$, is a solution to (1).

Here is a very important result: if Markov chain is irreducible, aperiodic and set $\mathcal{S}$ is finite then the stationary distribution $\vec{\pi}$ exists, it is unique and for any initial distribution $\vec{\mu}_0$ we have

$$\lim_{n \to \infty} \vec{\mu}_0 P^n = \vec{\pi}. \tag{2}$$

Definition 1.3. A continuous time stochastic process $\{X(t)\}_{t \geq 0}$ taking values in a countable state space S is a *Markov process* if and only if, for all $i, j \in S$, $s, t > 0$,

$$\begin{aligned}\mathbb{P}\{X(s+t) = j | \ X(s) = i \text{ and the history of the process prior to } s\} \\ = \mathbb{P}\{X(s+t) = j | \ X(s) = i\}.\end{aligned}$$

1.1.1. Harris positive recurrence

Consider a Markov chain X_n on a Polish space $\mathcal{S}$. Assume that there is a *recurrent set* R, i.e. for all $x \in \mathcal{S}$ hitting time $\tau_R = \inf\{n \geq 1 : X_n \in R\}$ is almost surely finite conditioned on $X_0 = x$. Assume that there exist an integer $l \geq 1$, probability measure Q on $\mathcal{S}$ and number $p \in (0,1)$ such that $p^{(l)}(x, \cdot) \geq pQ(\cdot)$, for all $x \in R$. We then say that Markov chain X_n possesses a *Harris property*, or that it is *Harris recurrent*.

Let X_0 be a random variable with distribution Q. Let $T_{R,k}$, $k \geq 1$ be the times at which the chain hits set R. Consider independent identically distributed random variables ζ_k, $k \geq 1$, taking values 0/1 and $\mathbb{P}\{\zeta_k = 1\} = p$. Let

$$K = \inf\{k \geq 1 : \ \zeta_k = 1\}. \tag{3}$$

If, in addition to Harris property, we also have $\mathbb{E}_Q T_{R,K} < \infty$, we then say that the chain is *Harris positive recurrent.*

There is extensive study on recurrence properties of Markov chains and its connections to convergence to stationary regime (see, e.g. Foss and Konstantopoulos (2004) for an overview). The same way there is a study of Markov processes, however, instead of analysing times τ_R the interest turns to $\tau_R^\varepsilon = \inf\{t \geq \varepsilon : X(t) \in R\}$.

§ 1.2. Cat-and-Mouse Markov chain and its appearances in practice

We analyse the dynamics of a stochastic process with dependent coordinates, commonly referred to as the Cat-and-Mouse (CM) Markov chain, and of its generalisations. Let $\mathcal{S}$ be a directed graph. Let $\{(C(n), M(n))\}_{n=0}^\infty$ denote the CM Markov chain on $\mathcal{S}^2$. Let $\{C(n)\}_{n=0}^\infty$, the location of the cat, be a Markov chain on $\mathcal{S}$ with transition matrix $P = (p(x,y)), x, y \in \mathcal{S})$ (here we treat S as the set of vertices, and everything else we may want to say about the graph structure is hidden on the matrix P). The second coordinate, the location of the mouse, $\{M(n)\}_{n=0}^\infty$ has the following dynamics:

- If $M(n) \neq C(n)$, then $M(n+1) = M(n)$,
- If $M(n) = C(n)$, then, conditionally on $M(n)$, the random variable $M(n+1)$ has distribution $(p(M(n), y), y \in \mathcal{S})$ and is independent of $C(n+1)$.

The process can be divided into cycles. In the beginning of each cycle the cat and the mouse are at the same location. Then they jump independently (the cat does not notice where the mouse jumps) and, if they happen to be at different locations, the mouse stays in hiding and not moving. When the cat hits the location of the mouse, a new cycle starts.

CM Markov chain is an example of models called Cat-and-Mouse games. CM games are common in game theory. As an example of practical use of this model we refer to Coppersmith *et al.* (1993). Here authors study the design and analysis of on-line algorithms.

An on-line algorithm instructs agents in the system to act only on the individually acquired information. On the other hand, an off-line algorithm can impose unified schedule and use global information. Even though the off-line algorithms bring the optimal results, the applications impose the developers to work with on-line algorithms.

A Cat-and-Mouse game is shown in comparison with a problem of synthesis of random walks on weighted graphs with positive real costs on their edges. The paper shows that a CM game is at the core of many on-line algorithms. Authors focus on two particular settings.

The first setting is the *k-server problem*, defined in Manasse *et al.* (1990). Here k mobile servers are considered on the vertices of a weighted graph under a sequence of requests for service. The requests are in the form of names of vertices, and each request is satisfied by placing a server at the requested vertex. The cost of satisfying a sequence of requests is the distance traversed by the servers. The goal is the development of on-line algorithms whose performance on any sequence of requests is as close as possible to the performance of the optimal off-line algorithm.

The second setting is the *metrical task system*, considered by Borodin *et al.* (1992). Here, an algorithm occupies one vertex of a weighted graph at any time. The cost of processing a given task depends on the state of the system. A schedule for a sequence $T^1, T^2, \ldots, T^k$ of tasks is a sequence $s_1, s_2, \ldots, s_k$ of states where s_j is the state where T^j is processed. The cost is a sum of processing and transition costs. Again, the goal is to construct an effective on-line algorithm which will handle the schedules. An interesting result discovered by Coppersmith *et al.* (1993) is that memoryless randomised algorithms can be competitive in various situations. A related model was studied by Baeza-Yates *et al.* (1993). The authors considered a problem of a robot searching for a certain object on a plane (a point or a line). The goal is to minimise the distance covered by this robot with various kinds of prior information (e.g., the distance to the object or the direction).

Let us come back to the memoryless case, i.e., to the case where both the cat and the mouse have fixed distributions of their jumps. In the monograph by Aldous and Fill (2002) the authors study the finite-time behaviour (e.g., hitting time of a set or mixing time) of various Markov chains and related objects on graphs. Chapter 3 of this monograph is dedicated to the study of reversible Markov chains, where the authors study the spectral representation and its consequences for the structure of hitting time distributions, the analogy with electrical networks, etc.. In this context the authors introduce two examples of CM games and discuss the average time between meetings of the cat and the mouse.

We are particularly interested in the paper of Litvak and Robert (2012). The paper introduces several settings with different graphs $\mathcal{S}$ (in our research we took interest in the case $\mathcal{S} = \mathbb{Z}$) and studies the limiting behaviour of the mouse. The authors analyse the connection between the Cat-and-Mouse Markov chain and the original on-line page-ranking algorithm from Abiteboul *et al.* (2003). Let us describe the on-line page-ranking algorithm mentioned above. Let a directed graph $\mathcal{S}$ describe the web: the nodes of the graph are the web pages and the html links between these pages are represented by the links of the graph. Let us model the browsing history of a customer by the process $\{C(n)\}_{n=0}^{\infty}$ with the transition matrix $P = (p(x, y)), x, y \in \mathcal{S})$ from the beginning of this section.

The goal is to find the page-rank of a node (a web page), i.e. to find the proportion of time the process C spends in that node or the stationary distribution of C.

Assume that at time 0 each node x has a weight $V_0(x) \geq 0$ such that $\sum_{x\in\mathcal{S}} V_0(x) = 1$. The process C updates the weights of the nodes it encounters on a random path through the graph. Let $h_0(x) = 0$ for $x \in \mathcal{S}$ and $h_n(x)$ be the approximation of the page-rank of a node x at time n. When the process C visits a node x at time n a following update happens:

- the weight $V_{n-1}(x)$ of node x is distributed to the neighbouring nodes according to transition probabilities $p(x, y)$,

$$V_n(x) = 0 \text{ and } V_n(y) = V_{n-1}(y) + V_{n-1}(x)p(x, y) \text{ for } y \neq x, \tag{4}$$

- the variable $h_n(x)$ accumulates the weight $V_{n-1}(x)$,

$$h_n(x) = h_{n-1}(x) + V_{n-1}(x) \text{ and } h_n(y) = h_{n-1}(y) \text{ for } y \neq x. \tag{5}$$

The estimate of the stationary distribution of C at x (the page-rank of x) is given by $\frac{h_n(x)}{\sum_{y\in\mathcal{S}} h_n(y)}$. It is shown by Litvak and Robert (2008) that this quantity converges to the stationary distribution of C as $n \to \infty$.

Let $(\mathcal{F}_n)$ denote the history of the motion of the cat, for $n \geq 0$, where $\mathcal{F}_n$ is the σ-field generated by the variables $\{C(k)\}_{k=0}^{n}$. Let $\mathbb{P}\{M(0) = x|\ C(0)\} = V_0(x)$ for $x \in \mathcal{S}$. Then Theorem 1 from Litvak and Robert (2012) gives the connection between the models:

$$(V_n(x), x \in \mathcal{S}) \overset{dist}{=} (\mathbb{P}\{M(n) = x |\ \mathcal{F}_{n-1}\}, x \in \mathcal{S}). \tag{6}$$

In particular, for $x \in \mathcal{S}$ we have $\mathbb{E}V_n(x) = \mathbb{P}\{M(n) = x\}$.

§ 1.3. Heavy-tailed distributions and related results

In this Subsection we give an overview of certain classes of distributions which are used extensively in Subsection **1.9**. We also present as series of results which are important for both this section and this manuscript in general.

Definition 1.4. A distribution F is called *(right) heavy-tailed* if

$$\mathbb{E}e^{\varepsilon\xi} = \infty$$

for any $\varepsilon > 0$, where ξ is a random variable having distribution F. If $\mathbb{E}e^{\varepsilon\xi} < \infty$ for some positive ε then the distribution F is called *light-tailed.*

1.3.1. Subexponential distributions

The following subclass of heavy-tailed distributions is of great importance in the theory.

Definition 1.5. A distribution F is called *long-tailed* if $\overline{F}(x) > 0$ for all x and

$$\mathbb{P}(\xi > x + t | \xi > x) = \frac{\overline{F}(x+t)}{\overline{F}(x)} \to 1 \tag{7}$$

as $x \to \infty$ for some (and, hence, for all) $t > 0$. Here ξ is a random variable with distribution F.

Intuitively, convergence (7) means that if (for a fixed t and for a large enough x) the value of the random variable exceeds the level x then it also exceeds the level $x + t$ with probability close to 1.

Definition 1.6. A distribution F on $\mathbb{R}^+ = [0, \infty)$ is called *subexponential* if $\overline{F}(x) > 0$ for all x and

$$\frac{\mathbb{P}(\xi_1 + \xi_2 > x)}{\mathbb{P}(\xi_1 > x)} = \frac{\overline{F^{*2}}(x)}{\overline{F}(x)} \to 2 \tag{8}$$

as $x \to \infty$ for independent random variables ξ_1 and ξ_2 both having distribution F.

A distribution F on $\mathbb{R}$ is called *subexponential* if the distribution F^+ given by its distribution function $F^+(x) = F(x)\,\mathbb{I}(x \geq 0)$ is subexponential.

In Embrechts and Goldie (1982) it is proved that the property (8) of subexponential distributions can be generalised to the sum of an arbitrary number of independent identically distributed random variables. Namely, the following characterisation of subexponential distributions is valid.

Proposition 1.7. *Let $\{\xi_i\}_{i=1}^{\infty}$ be a sequence of independent random variables with a distribution F. The distribution F is subexponential if and only if*

$$\mathbb{P}(\xi_1 + \ldots + \xi_n > x) \sim n\mathbb{P}(\xi_1 > x)$$

or, equivalently, if and only if

$$\mathbb{P}(\xi_1 + \ldots + \xi_n > x) \sim \mathbb{P}(\max\{\xi_1, \ldots, \xi_n\} > x) \tag{9}$$

as $x \to \infty$ for some $n \geq 2$.

Throughout this thesis, we say that $a(x) \sim b(x)$ as $x \to \infty$ if $\dfrac{a(x)}{b(x)} \to 1$ as $x \to \infty$.

The relation (9) may be interpreted in the following way: with probability close to 1, the sum of subexponential random variables exceeds a sufficiently large level x due to the fact that one of the summands exceeds this level.

In Subsection **1.3.2** we show several results concerning the generalisation of the property (8) to the sum of random variables over an independent random time.

We now mention some properties of subexponential random variables that will be used in our work.

Proposition 1.8. *Let $\eta_1, \eta_2, \ldots, \eta_n$ be independent random variables and let F be a subexponential distribution such that for every $i = 1, 2, \ldots, n$ the asymptotic equivalence $\mathbb{P}(\eta_i > x) \sim c_i\overline{F}(x)$ as $x \to \infty$ holds with some constant $c_i > 0$. Then*

$$\mathbb{P}(\eta_1 + \eta_2 + \ldots + \eta_n > x) \sim (c_1 + c_2 + \ldots + c_n)\overline{F}(x).$$

Proposition 1.9. *Assume that the tails of the distributions F and H are asymptotically equivalent, i.e. $\overline{H}(x) \sim c\overline{F}(x)$ for some $c > 0$. If F is subexponential, then H is subexponential too.*

One can find the proof of Proposition **1.8** in Cline (1986), Proposition **1.9** — in Embrechts and Goldie (1982).

We remark that the class of subexponential distributions is not closed under the operation of convolution: if random variables ξ and η are subexponential, their sum $\xi + \eta$ is not always subexponential (see Leslie (1989)).

Here are some popular examples of subexponential distributions:

The Pareto distribution with the tail $\overline{F}(x) = \left(\frac{\kappa}{\kappa + x}\right)^{\beta}$ where $\beta, \kappa > 0$;

The log-normal distribution with the density $\frac{e^{-(\ln x - \ln \beta)^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}x}$ where $\beta > 0$;

The Weibull distribution with the tail $\overline{F}(x) = e^{-x^{\beta}}$ where $\beta \in (0, 1)$.

We introduce another important subclass of subexponential distributions, which we use throughout the thesis.

Definition 1.10. A distribution F is called *regularly varying* (at infinity) with parameter β if its tail is given by

$$\overline{F}(x) = x^{-\beta} l(x)$$

where $l(x) > 0$ is *a slowly varying function*, i.e. such that

$$\frac{l(tx)}{l(x)} \to 1$$

as $x \to \infty$ for any $t > 0$.

The definition and various properties of the regularly varying distributions may be found in Bingham *et al.* (1987). It was proved in Section VIII.8 of Feller (1971b) that all such distributions are subexponential.

1.3.2. Tail asymptotics for randomly stopped sums

In this Subsection we present some useful results concerning randomly stopped sums which we use in Sections **3** and **4**.

Let $\xi_1, \xi_2, \ldots$ be positive i.i.d. r.v.'s with a common distribution function F. Let $S_0 = 0$ and $S_k = \xi_1 + \ldots \xi_k$, $k \geq 1$. Let τ be a counting r.v. with distribution function G, independent of $\{\xi_k\}_{k=1}^{\infty}$. For a general overview concerning asymptotics of tail-distribution of S_τ see, e.g., Denisov *et al.* (2010) and references therein. The key result in the theory of subexponential distributions is the following: if F

is subexponential and G is light-tailed, then

$$\mathbb{P}\{S_\tau > n\} \sim \mathbb{E}\tau \overline{F}(n), \text{ as } n \to \infty. \tag{10}$$

For our purposes we need distribution G to be more general, when ξ_1 has an infinite mean. The next result follows from Corollary **3** from Foss & Zachary (2003).

Proposition 1.11. *Assume that $\overline{F}(x) \sim l_1(x)/x^\alpha$, $\alpha \in [0,1)$ and τ has any distribution with $\mathbb{E}\tau < \infty$. Then*

$$\mathbb{P}\{S_\tau > n\} \sim \mathbb{E}\tau \mathbb{P}\{\xi > n\} \textit{ as } n \to \infty. \tag{11}$$

The next result is very useful for us in Section **4**. It was proved via Tauberian theorems (see Appendix A.1).

Proposition 1.12. *Assume that $\overline{F}(x) \sim l_1(x)/x^\alpha$ and $\overline{G}(x) \sim l_2(x)/x^\beta$, $\alpha, \beta \in (0,1)$. Then*

$$\mathbb{P}\{S_\tau > n\} \sim n^{-\alpha\beta} \frac{\Gamma^\beta(1-\alpha)\Gamma(1-\beta)}{\Gamma(1-\alpha\beta)} l_1^\beta(n)\, l_2\left(\frac{n^\alpha}{\Gamma(1-\alpha) l_1(n)}\right), \textit{ as } n \to \infty. \tag{12}$$

1.3.3. Hitting times for random walks and renewal processes

In this Subsection we discuss random walks on integer lattice $\mathbb{Z}$. Let $\{\xi_n\}_{n=1}^\infty$ be independent identically distributed random variables. Let $S_0 = 0$ and $S_n = \sum_{k=1}^n \xi_k$, for $n \geq 1$. Denote $\tau_x = \inf\{n \geq 1 : S_n = x\}$, for $x \in \mathbb{Z}$.

Let ξ_1 take values ± 1 with equal probabilities $1/2$. Then the random walk possesses a 'continuity property', meaning there is no jumps over any point x. This, for example, leads to the following property: take τ' and τ'' two independent copies of τ_1 and one can write $\tau_2 \stackrel{d}{=} \tau' + \tau''$. We have a very nice explicit result with regards to the tail-asymptotics of τ_0 (see, e.g., Section III.2 in Feller (1971) for related result):

$$\mathbb{P}\{\tau_0 > n\} \sim \sqrt{\frac{2}{\pi n}}, \text{ as } n \to \infty. \tag{13}$$

Due to the symmetry of the system, we have $\tau_0 \stackrel{d}{=} 1 + \tau_1$. Therefore, we get explicit tail-asymptotics for every τ_x.

Now, let us assume that ξ_1 has a general aperiodic distribution with zero-mean and finite variance σ^2. Denote

$$a^*(x) = \mathbb{I}(x = 0) + \sum_{n=0}^{\infty} \left(\mathbb{P}\{S_n = 0\} - \mathbb{P}\{S_n = -x\}\right). \tag{14}$$

Then we have the following result (see, e.g., Chapter VII of Spitzer (1964)):

$$\mathbb{P}\{\tau_x > n\} \sim \sigma\sqrt{\frac{2}{\pi n}} a^*(x), \text{ as } n \to \infty. \tag{15}$$

Even though the function $a^*(x)$ is not explicit there is extensive knowledge on its asymptotics. Nevertheless, it was not sufficiently detailed in order to prove our results and we used the analysis of the characteristic function of τ_x from Uchiyama (2011a) to acquire our results.

Another known result we would like to mention concerns the two-dimensional integer lattice $\mathbb{Z}^2$. If ξ_1 take values on unit vectors with equal probabilities 1/4, then

$$\mathbb{P}\{\tau_0 > n\} \sim \frac{\pi}{\ln n}, \text{ as } n \to \infty \tag{16}$$

(see, e.g., Chapter III of Spitzer (1964)). Thus, in this case the tail-distribution is a slowly varying function, which brings certain complications in comparison with the one-dimensional case.

In the Cat-and-Mouse model the mouse waits for the moment when the cat reaches a certain position. The waiting time closely relates to τ_x. The analysis of the trajectories of the mouse is roughly equivalent to the analysis of the number of the meetings time-instants. Let $\{\zeta_n\}_{n=1}^{\infty}$ be independent identically distributed counting random variables. Let $S_0^\tau = 0$ and $S_n^\tau = \sum_{k=1}^{n} \zeta_k$, for $n \geq 1$. Define a *renewal process*

$$\nu(t) = \max\{n \geq 1 : \ S_n^\tau \leq t\}, \ t \geq 0. \tag{17}$$

The following results can found in Section XI.5 in Feller (1971b). If $\mathbb{E}\zeta_1 = \mu$ and $\mathbf{Var}\zeta_1 = \sigma^2$ then

$$\frac{\nu_t - t\mu^{-1}}{\sqrt{t\sigma^2\mu^{-3}}} \Rightarrow \psi \sim N(0,1), \text{ as } t \to \infty. \tag{18}$$

However, in our models the time-intervals have an infinite mean. Nevertheless, similar to (15), they have regularly varying tail-distributions. If $\mathbb{P}\{\zeta_1 > x\}$ is regularly varying with parameter $\alpha \in (0,1)$, there exists a random variable φ such that

$$\mathbb{P}\{\zeta_1 > t\}\nu(t) \Rightarrow \varphi, \text{ as } t \to \infty. \tag{19}$$

§ 1.4. Lévy processes, Brownian motion and local time

D e f i n i t i o n 1.13. A random process $\{X_t\}_{t\geq 0}$ is called a *Lévy process* if

(a) for any $s, t \geq 0$ the increment $X_{t+s} - X_t$ does not depend on the process $\{X_v\}_{0\leq v\leq t}$ and has the same distribution as X_s;

(b) trajectories of this process almost surely belong to class $D([0,\infty))$, i.e. are right-continuous for $t \geq 0$ and have left limits for $t > 0$.

A very important example of a Lévy process is a *Brownian motion* (or a *Wiener process*), which is Lévy process $\{B_t\}_{t\geq 0}$ such that B_1 has a standard normal distribution. To formulate some upcoming results we need to define *local time* of Brownian motion. There are multiple equivalent definitions for this object in the literature.

D e f i n i t i o n 1.14. A local time $L_x(t)$ of a Brownian motion B_t at point x up to time t is

$$L_x(t) = \lim_{\varepsilon\to 0} \frac{1}{\varepsilon} \int_0^t \mathbb{I}(B_s \in [x, x+\varepsilon])ds = \lim_{\varepsilon\to 0} \frac{1}{2\varepsilon} \int_0^t \mathbb{I}(B_s \in [x-\varepsilon, x+\varepsilon])ds \tag{20}$$

§ 1.5. Weak convergence for processes taking values in $D[[0,\infty), \mathbb{R}]$

Most of our main results relate to the weak convergence of stochastic processes. Thus, in this subsection we recall definitions of the $\mathcal{J}_1$-topology (see, e.g., Skorokhod (1956)) and make necessary comments for better understanding of our results and proofs.

Let $D[[0,T],\mathbb{R}]$ denote the space of all right continuous functions on $[0,T]$ having left limits. For any $g \in D[[0,T],\mathbb{R}]$, let $\|g\| = \sup_{t\in[0,T]} |g(t)|$. Let Λ be the set of increasing continuous functions $\lambda : [0,T] \to [0,T]$, such that $\lambda(0) = 0$ and $\lambda(T) = T$. Let λ_{id} denote the identity function. Then

$$d_{\mathcal{J}_1,T}(g_1, g_2) = \inf_{\lambda\in\Lambda} \max(\|g_1 \circ \lambda - g_2\|, \|\lambda - \lambda_{id}\|) \tag{21}$$

defines a metric inducing $\mathcal{J}_1$.

It is important to remember that the main reason such metrics were introduced is to deal with the convergence of discontinuous processes. However, in cases where the limiting function is continuous the analysis of convergence in $d_{\mathcal{J}_1,T}$-metric is equivalent to the analysis in uniformal metric.

On the space $D[[0,\infty),\mathbb{R}]$ the $\mathcal{J}_1$-topology is defined by the metric

$$d_{\mathcal{J}_1,\infty}(g_1,g_2) = \int_0^\infty e^{-t}\min(1, d_{\mathcal{J}_1,t}(g_1,g_2))dt. \tag{22}$$

Convergence $g_n \to g$ in $(D[[0,\infty),\mathbb{R}],\mathcal{J}_1)$ means that $d_{\mathcal{J}_1,T}(g_n,g) \to 0$ for every continuity point T of g.

Definition 1.15. Let $\{\{X_n(t)\}_{t\geq 0}\}_{n=1}^\infty$ and $\{X(t)\}_{t\geq 0}$ be stochastic processes with trajectories from $D[[0,\infty),\mathbb{R}]$. We say that weak convergence

$$\{X_n(t),\ t\geq 0\} \overset{\mathcal{J}_1}{\Rightarrow} \{X(t),\ t\geq 0\}, \tag{23}$$

holds if

$$\mathbb{E}f(\{X_n(t)\}_{t\geq 0}) \to \mathbb{E}f(\{X(t)\}_{t\geq 0}),\ \text{as } n\to\infty, \tag{24}$$

for any continuous and bounded function f on $D[[0,\infty),\mathbb{R}]$ endowed with $\mathcal{J}_1$-topology.

In Section 2.6 of Skorokhod (1956), for processes from $D[[0,1],\mathbb{R}]$ we can find that necessary and sufficient conditions for $X_n(t) \overset{\mathcal{J}_1}{\Rightarrow} X(t)$ are that

- $X_n(t)$ converges to $X(t)$ on an everywhere dense set containing 0 and 1, and
- convergence $\lim_{c\to 0}\overline{\lim}_{n\to\infty}\Delta_{\mathcal{J}_1}(c, X_n(t)) = 0$ holds, where

$$\Delta_{\mathcal{J}_1}(c, y(t)) = \sup_{t-c<t_1<t\leq t_2<t+c}\min(\|y(t_1)-y(t)\|, \|y(t)-y(t_2)\|) \tag{25}$$

Additionally we introduce the following well-known result (see, e.g., Billingsley (1968)) which we use later on in our analysis.

Proposition 1.16. *Let $\{X_n\}_{n=1}^\infty$ and $\{Y_n\}_{n=1}^\infty$ be two sequences of stochastic processes with trajectories from $D[[0,\infty),\mathbb{R}]$. Given $d_{\mathcal{J}_1,\infty}(X_n,Y_n) \overset{a.s.}{\to} 0$, we have*

$$X_n - Y_n \overset{\mathcal{J}_1}{\Rightarrow} 0. \tag{26}$$

As mentioned above, if the limiting process is continuous then the metrics become roughly equivalent. In the thesis we treat only the cases where the limit is almost sure continuous. Thus, we omit the J_1-topology and use the standard notation

$$\{X_n(t),\ t \geq 0\} \overset{\mathcal{D}}{\Rightarrow} \{X(t),\ t \geq 0\}. \tag{27}$$

§ 1.6. Scaling results for the mouse on $\mathbb{Z}, \mathbb{Z}^2$ or $\mathbb{Z}^+$

As mentioned above, we are particularly interested in the paper of Litvak and Robert (2012). Here we state several results of this paper, which gave the main motivation for our research. The paper describes scaling properties of the (non-Markovian) sequence $\{M(n)\}_{n=0}^{\infty}$ for a specific transition matrix P when $\mathcal{S}$ is either $\mathbb{Z}, \mathbb{Z}^2$ or $\mathbb{Z}^+$. Further, the authors consider a general recurrent Markov chain $\{C(n)\}_{n=0}^{\infty}$ and provide recurrence properties of the Markov chain $\{(C(n), M(n))\}_{n=0}^{\infty}$ (we omit the latter result).

Let the components live on the integer line $\mathbb{Z}$ ($\mathcal{S} = \mathbb{Z}$) and take the transition matrix P such that $p(x, x+1) = p(x, x-1) = 1/2$. It was proven in Theorem 3 of Litvak and Robert (2012) that the convergence in distribution

$$\left\{ \frac{1}{\sqrt[4]{n}} M([nt]),\ t \geq 0 \right\} \overset{\mathcal{D}}{\Rightarrow} \{B_1(L_{B_2}(t)),\ t \geq 0\},\ \text{as } n \to \infty \tag{28}$$

holds, where $B_1(t)$ and $B_2(t)$ are independent standard Brownian motions on $\mathbb{R}$ and $L_{B_2}(t)$ is the local time process of $B_2(t)$ at 0.

This result looks natural, since the mouse, observed at the meeting times with the cat, is a simple random walk. From the Central Limit Theorem, we can loosely say that in n steps the distance between a random walk and the origin is of order $n^{1/2}$. The time intervals between meeting time-instants are independent and identically distributed. They have the same distribution as the time needed for the cat (also a simple random walk) to get from 1 to 0, which has a regularly varying tail with parameter 1/2 (see, e.g., Spitzer (1964)). Again, we can loosely say that up to time n the number of meeting time-instants is of order $n^{1/2}$. Thus, after n steps of the system we expect the distance between the mouse and the origin to be of order $\sqrt[4]{n} = (n^{1/2})^{1/2}$. Local time $L_{B_2}(t)$ can be interpreted as the scaled duration of time the cat and the mouse spend together.

Now, let $\mathcal{S} = \mathbb{Z}^2$ and let transition matrix P satisfy

$$p(x, x + (1, 0)) = p(x, x + (0, 1)) = p(x, x - (1, 0) = p(x, x - (0, 1))) = \frac{1}{4}. \tag{29}$$

Let $W(t) = (W_1(t), W_2(t))$ and $B(t)$ be independent two-dimensional and one-dimensional Brownian motions. Denote $L_B(t)$ as the local time process of $B(t)$ at 0. For $t \geq 0$ let $T_t = \inf\{s \geq 0 : \ B(s) = t\}$. Theorem 4 from Litvak and Robert (2012) proves that for a positive constant C_R (the paper describes the explicit formula for this constant) the following convergence in distribution holds:

$$\left\{\frac{1}{\sqrt{n}} M([e^{nt}]), \ t \geq 0\right\} \overset{\mathcal{D}}{\Rightarrow} \{W(C_R L_B(T_t)), \ t \geq 0\}, \text{ as } n \to \infty. \tag{30}$$

The main difference of this case is that the time intervals between meeting time-instants have a distribution with a tail of order $\log n$ (which means it is slowly varying or regularly varying with index 0). This introduces a great challenge to acquire the scaling for $M([nt])$. Nevertheless, after e^n steps of the system we expect the number of meeting time-instants to be of order n, which gives us an intuition for this result.

Finally, let $\mathcal{S} = \mathbb{Z}^+$ and let transition matrix P satisfy

$$\begin{cases} p(x, x+1) = p, & x \geq 0, \\ p(x, x-1) = 1 - p, & x \neq 0, \\ p(0, 0) = 1 - p. \end{cases} \tag{31}$$

Take $p \in (0, 1/2)$ and let $\varrho = p/(1-p)$. We assume that at time zero the mouse is far away from the cat. Define

$$s_1 = \inf\{n \geq 0 : \ C(n) = M(n)\} \text{ and } t_1 = \inf\{n \geq s_1 : \ C(n) = 0\} \tag{32}$$

and, for $k \geq 1$,

$$s_{k+1} = \inf\{n \geq t_k : \ C(n) = M(n)\} \text{ and } t_1 = \inf\{n \geq s_{k+1} : \ C(n) = 0\}. \tag{33}$$

Let $\mathbb{I}(A)$ denote the indicator function, meaning that for event A

$$\mathbb{I}(A) = \begin{cases} 1, & \text{if } A \text{ holds} \\ 0, & \text{otherwise.} \end{cases} \tag{34}$$

Theorem 5 from Litvak and Robert (2012) states that if $M(0) = n$ and $C(0) = n$ then for a positive random variable W (again, the paper states the explicit formula) the following convergence in distribution holds:

$$\left\{ \frac{M([t\varrho^{-n}])}{n} \mathbb{I}[t < \varrho^n t_n],\ t \geq 0 \right\} \overset{\mathcal{D}}{\Rightarrow} \{\mathbb{I}[t < W],\ t \geq 0\}, \text{ as } n \to \infty. \tag{35}$$

In this case, when the mouse is at n 'far away from the origin' the order of time needed for the cat to reach the same level is ϱ^{-n}. After certain amount of time the cat will return to zero and the procedure restarts. Each time, the mouse will be pushed to zero by the drift, until eventually it will stay in a neighbourhood of zero.

§ 1.7. Stable laws

In this subsection we discuss *stable laws* and related limiting results. For reference see, e.g., Section 6.1 in Feller (1971b). Assume that X and $\{X_k\}_{k=1}^{\infty}$ are mutually independent identically distributed random variables. Let $S_n = \sum_{k=1}^{n} X_k$.

Definition 1.17. Let R be the distribution of X. Distribution R is stable (in the broad sense) if for each n there exists $c_n > 0$, γ_n such that

$$S_n \overset{d}{=} c_n X + \gamma_n, \tag{36}$$

and R is not concentrated at one point. It is stable in the strict sense if the above holds with $\gamma_n = 0$ for $n > 1$.

The normal distribution is of course a very important example of a stable distribution. Now, we formulate a property which is analogous to the convergence in the Central Limit Theorem.

Definition 1.18. The distribution F of the independent random variables X_k belongs to the domain of attraction of a distribution R if there exists norming constants $b_n > 0$, d_n such that the distribution of $b_n^{-1}(S_n - d_n)$ tends to R.

Thus, the Central Limit Theorem states that if $\sigma^2 = \mathbf{Var}X_1 < \infty$ then F belongs to the domain of attraction of standard normal distribution with $b_n = \sigma\sqrt{n}$ and $d_n = n\mathbb{E}X_1$. It is important that distribution R possesses a domain of attraction if and only if it is stable.

We need a related result concerning regularly varying tails. Denote

$$F_+(x) = \mathbb{P}\{X_1 \geq x\},\ F_-(x) = \mathbb{P}\{X_1 \leq -x\} \text{ and } F_\pm(x) = F_+(x) + F_-(x). \quad (37)$$

Assume that $F_\pm$ is regularly varying at infinity with parameter $\alpha \in (0, 2]$. Additionally, there exists a limit

$$\lim_{x\to\infty} \frac{F_+(x)}{F_\pm(x)} = \varrho_+ \in [0, 1]. \quad (38)$$

If $\mathbb{E}|X_1| < \infty$ assume that the mean is zero. Then for $\alpha \neq 1$ distribution F belongs to the domain of attraction of a stable distribution with $b_n = F_\pm^{(-1)}(1/n)$ (which is regularly varying with parameter $-1/\alpha$) and $d_n = 0$ (see, e.g., Section 1.5 of Borovkov and Borovkov (2008)). For the case $\alpha = 1$, the centering factor d_n needs additional discussion. In the thesis we consider only the cases when $d_n = 0$.

The next important result comes from Example 11.2.18 of Meerschaert and Scheffler (2001). Let $((X_n^{(1)}, \ldots, X_n^{(d)}))_{n=1}^\infty$ is a sequence of independent identically distributed random vectors. Let $((b_n^{(i)})_{i=1}^d)_{n=1}^\infty$ be regularly varying sequences. Assume that for a d-dimensional random vector A with stable distribution holds

$$\left(\frac{\sum_{k=1}^n X_k^{(1)}}{b_n^{(1)}}, \ \ldots, \ \frac{\sum_{k=1}^n X_k^{(d)}}{b_n^{(d)}} \right) \Rightarrow A, \text{ as } n \to \infty. \quad (39)$$

Then there is a Lévy process $A(t)$ such that

$$\left(\frac{\sum_{k=1}^{[ct]} X_k^{(1)}}{b_{[ct]}^{(1)}}, \ \ldots, \ \frac{\sum_{k=1}^{[ct]} X_k^{(d)}}{b_{[ct]}^{(d)}} \right) \overset{FD}{\Rightarrow} A(t), \text{ as } c \to \infty, \quad (40)$$

where $\overset{FD}{\Rightarrow}$ means weak convergence of all finite-dimensional marginal distributions, i.e., for every finite $N \geq 1$ and $0 \leq t_1 \leq \ldots t_N$ there is a joint weak convergence. Thus, for such random vector A with stable distribution we have a Lévy process $A(t)$ such that $A(1) \overset{d}{=} A$ and we call it a *Lévy process generated by A*.

§ 1.8. General overview of our results

In the Thesis we restrict ourselves to the case $\mathcal{S} = \mathbb{Z}$. However, we do not necessarily restrict the jumps of the mouse to have the same distribution as the jumps of the cat. We provide two generalisations of the CM Markov chain introduced above.

The first generalisation relates to the jump distribution of the components. First, we assume that the cat continues to follow a simple random walk while, given $C(n) = M(n)$, the random variable $M(n+1) - M(n)$ has a general distribution which has a finite first moment and belongs to the strict domain of attraction of a stable law with an index $\alpha \in [1, 2]$, with a normalising function $\{b(n)\}_{n=1}^{\infty}$ (note that distributions with a finite second moment belong to the domain of attraction of a normal distribution). It is important to notice that the cases of zero and non-zero first moment are different and leads to different scalings. Nevertheless, it will not influence the proofs that much. For the case of zero mean we find a weak limit of $\{b^{-1}(\sqrt{n})M([nt])\}_{t\geq 0}$ as $n \to \infty$. This model is more challenging than the classical setting because, when the mouse jumps, the value of this jump and the time until the next jump may be dependent. Also, if the jump distribution of the mouse has an infinite second moment we can not use classical results such as Theorem 5.1 from Kasahara (1984). Next, we consider the case where both components have general distributions with finite second moments. Here our results take into account the approach developed by Uchiyama (2011a).

In the second generalisation we add more components (we will refer to the objects whose dynamics these components describe as "agents") to the system, with keeping the chain "hierarchy". For instance, adding one extra agent (we refer to it as the dog), acting on the cat the same way as the cat acts on the mouse, slows down the cat and, therefore, also the mouse. We are interested in the effect of this on the scaling properties of the process. Recursive addition of further agents will slow down the mouse further. For the system with three agents we find a weak limit of $\{n^{-1/8}M([nt])\}_{t\geq 0}$ as $n \to \infty$. The system regenerates when all the agents are at the same point. Therefore, if we find the tail asymptotics of the time intervals between this events, we can split the process into independent identically distributed cycles and use classical results (for example, Kasahara (1984)).

For the systems with an arbitrary finite number of agents, we provide a relatively simple result on the weak convergence, for fixed $t > 0$. In this case, the path analysis becomes quite difficult and we have not yet found the asymptotics of the time intervals between regeneration points. Nevertheless, we transform the number of jumps for any agent and use the induction and a result from Dobrushin (1955), concerning the limiting behaviour of compound processes, to prove a weak convergence for each individual component.

§ 1.9. Related results

There are many related models in applied probability where time evolution of the process may be represented as a multi-component Markov chain where one of the components has independent dynamics and forms a Markov chain itself. Papers by Gamarnik (2004) and Gamarnik and Squillante (2005) consider bin packing process. Let N be a (fixed) positive integer. One bin of size N arrives at discrete times $t = 0, 1, 2, \ldots$. Items arrive at discrete times according to a general stochastic process. Item sizes take on values $1, 2, \ldots, N$. The problem is to fill the bin with items in possession in time-interval $[t, t+1)$ and prevent overstocking of the items. The items form queues $Q_i(t)$ where the index i refers to the size of an item. Thus, the system can be represented by the process $(Q_N(t), \ldots, Q_1(t))$. The authors consider the *largest first* policy: at each time we try to fit in the largest possible item. Thus, process $Q_N(t)$ is independent of other queues (this often leads to $Q_N(t)$ being a discrete Markov process). Queue $Q_{N-1}(t)$ can decrease only if $Q_N(t) = 0$. In the Cat-and-Mouse Markov chain a similar connection can be seen for $(C(n) - M(n), M(n))$: process $M(n)$ can change only if $C(n) - M(n) = 0$. The authors analyse the stability conditions for such model via combination of a Lyapunov function technique and matrix-analytic methods.

A separate example of the aforementioned relationship between coordinates can be found in Borst *et al.* (2008). The authors consider a multidimensional birth-and-death process where the birth-rates are fixed and death-rates may depend on the current state of the whole system. The stochastic monotonicity of the processes is used to reduce the analysis to the case where a subset of coordinates forms a Markov process and the rest depends on it through death-rates. The authors derive necessary and sufficient conditions for the stability of the system.

Typically the described dependences are modelled using Markov modulation. Consider a regenerative process $X = \{X_n,\ n \geq 1\}$ with state space $\mathcal{X}$. A random walk $\{S_n,\ n \geq 0\}$, defined by $S_0 = 0$ and $S_n = \xi_1 + \ldots + \xi_n$ for $n \geq 1$, is modulated by the process X if

- conditionally on X, the random variables ξ_n, $n \geq 1$, are independent,

- for some family $\{F_x,\ x \in \mathcal{X}\}$ of distribution functions such that, for each y, $F_x(y)$ is a measurable function of x, we have

$$\mathbb{P}\{\xi_n \le y|\ X\} = \mathbb{P}\{\xi_n \le y|\ X_n\} = F_{X_n}(y), \text{ a.s.} \tag{41}$$

For the Cat-and-Mouse Markov chain we can take process X as $X_n = C(n) - M(n)$ and the increments in the form $\xi_n \mathbb{I}(C(n-1) = M(n-1))$.

There is a number of papers considering the large deviations problems for such models. Assume that $\mathcal{X}$ is finite. The papers by Arndt (1980) and Alsmeyer and Sgibnev (1999) consider a certain class of tail distributions for increments ξ_n (including subexponential distributions with finite mean). Process X is an irreducible aperiodic Markov chain. Denote

$$M_\infty = \sup_{n \ge 0} S_n \text{ and } \eta(y) = \inf\{n \ge 1 :\ S_n > y\}. \tag{42}$$

The authors study the limiting behaviour of

$$W_{ij}(y) = \mathbb{P}\{M_\infty > y,\ X_{\eta(y)} = j|\ X_0 = i\}. \tag{43}$$

A comparison between the conditions on the increments can be found in Section 5 of Alsmeyer and Sgibnev (1999).

A similar setting occur in Jelenkovic and Lazar (1998). The authors study the tail asymptotic of M_∞ for subexponential increments via Wiener-Hopf factorisation (see also Asmussen (1989, 1991)). Additionally, the authors apply their random walk result to investigate two canonical queueing scenarios that are of practical importance in engineering broadband multiplexers. One of these scenarios considers a modulated queueing system, where the process X changes its values according to a subexponential renewal process. The absence of exponential moments is of interest to us, as in the Cat-and-Mouse Markov chain the time-intervals between meetings have infinite mean.

In the paper by Wang and Liu (2011) the authors discuss a different problem for a similar model. They assume that the increments have a long-tailed distribution. Then the authors obtain necessary and sufficient conditions for $\mathbb{P}\{M_\infty \in (x, x+z]\}$ to have an asymptotic form, as $x \to \infty$.

Foss *et al.* (2007) consider a more general setting. The state space $\mathcal{X}$ for process X does not need to be finite, but it is assumed that X to be positive recurrent.

The authors consider cases of both discrete and continuous time. Assuming the increments have a negative drift and are heavy-tailed the authors present natural conditions under which the tail distribution of M_∞ can be computed. A key instrument of the proofs is the so-called "principle of a single big jump", and the overall analysis is entirely probabilistic.

Similar questions for a different setup (although related the queueing examples of Jelenkovic and Lazar (1998)) are considered in Hansen and Jensen (2005). The authors consider a Markov-modulated reflected additive process, i.e. instead of random walk S_n we have a process $W_n = (W_{n-1} + \xi_n)^+$. Denote a stopping time $\sigma = \inf\{n \geq 1 : \ W_n = 0, \ X_n = X_0\}$. For $\mathcal{M}_\sigma = \max_{0\leq n\leq\sigma} W_n$ the authors find tail asymptotics. An interesting feature of the heavy-tailed framework is that the distribution of the cycle maximum and the invariant distribution of W_n are not in general tail equivalent in contrast to the light-tailed case. Thus, the asymptotic behaviour of the running maximum of W cannot be related directly to the tail of the invariant distribution. Nevertheless, the regenerative structure of the process allows to derive the aforementioned behaviour on the basis of tail asymptotics of $\mathcal{M}_\sigma$.

Asmussen *et al.* (1994) investigate Markov-modulated risk processes in the presence of heavy tails (see also Asmussen (1991), Lu and Li (2005), Asmussen and Foss (2014) and references therein). The authors consider a case where X_t is a continuous-time Markov process and S_t linearly decreases unless there is an arrival (which depends on X_t). The paper describes the asymptotic behaviour of $\psi(u) = \mathbb{P}\{M_\infty > u\}$, as $u \to \infty$. It is interesting that, even though in the light-tailed case Markov modulation changes the order of magnitude of the ruin probabilities for large u, in the heavy-tailed scenario the order of magnitude remains the same.

We now discuss further stability problems stability problems. Foss *et al.* (2012) consider a general Markov-modulated Markov chain (X_t, Y_t). Process X_t is considered to be Harris-ergodic. Process Y_t is multidimensional and it satisfies certain Foster-type conditions. Under these and some further natural assumptions the authors find a positive recurrent set for the system. Then the authors study the stability of two systems with multiple access random protocols in a changing environment.

Foss *et al.* (2018) consider *stochastic recursive sequences* $Z_{t+1} = f(Z_t, X_t)$ with a monotone increasing function f and *regenerative* sequence X_t. Under certain mixing conditions the authors prove stability. The results are then applied to three examples of growth, savings, and risk-sharing models.

Models where both components' dynamics depend on each other are considerably more difficult. Shah and Shin (2012) consider two representative examples of queueing network models that arise in the context of emerging communication networks. The first model captures a randomly varying number of packets travelling through a collection of wireless nodes communicating through a shared medium. The second model is a buffered circuit switched network capturing the randomness in calls. The authors introduce a scheduling algorithm for these models. Using a proper time-scale, the authors show a distinction between the network queueing dynamics and the scheduling dynamics induced by the algorithm. Essentially, the setup is transformed into a model where one of the components behaves as "almost stable", changing only relatively slowly compared to another component. Then the authors exhibit a proper Lyapunov functions and prove stability.

Georgiou and Wade (2014) consider Markov chains (Z_n, X_n) on $\mathbb{Z}^+ \times \mathcal{X}$ where $\mathcal{X}$ is finite. Neither coordinate is assumed to be Markov. It is assumed that the increments of Z_n satisfy certain moment conditions, and that, roughly speaking, process X_n is "almost Markov" when the first coordinate is large. The latter allows the authors to probe the recurrence phase transitions. The paper exhibit a recurrence classification in terms of increment moment parameters of Z_n and the stationary distribution for the large-Z limit of X_n.

§ 1.10. Structure of Part I

The rest of the first Part of the Thesis is structured as follows. In Subsection **1.11** we give a slightly different way to model the 'standard' CM Markov chain and reiterate the known result. Then in Subsection **1.12** we define its generalisation onto a general jump distribution of the mouse and formulate the first result on its limiting behaviour. Then we formulate a result in the case where both components have general distributions with finite second moments. In Subsection **1.13** we introduce another extension of the model, a Dog-and-Cat-Mouse (DCM) process, and formulate a corresponding limiting result. In Subsection **1.14** we consider a further extension of the DCM model onto N agents and formulate a corresponding result.

Our proofs of the formulated results are given in Sections **2** to **4**. Section **2** deals with hierarchical model of arbitrary length N and contains the proof of the result of Subsection **1.14**. Section **3** analyses the CM model with a general distribution of the mouse's jump and contains the proof the results of Subsection **1.12**. Finally, Section **4** considers the DCM model and includes the proof of the result of Subsection **1.13**.

§ 1.11. "Standard" Cat-and-Mouse Markov chain on $\mathbb{Z}$ $(C \to M)$

Let ξ take values ± 1 with equal probabilities $1/2$. Let $\{\xi_n^{(1)}\}_{n=1}^{\infty}$ and $\{\xi_n^{(2)}\}_{n=1}^{\infty}$ be two mutually independent sequences of independent copies of ξ. We define the dynamics of CM Markov chain $(C(n), M(n))_{n=0}^{\infty}$ as follows:

$$C(n) = C(n-1) + \xi_n^{(1)}, \tag{44}$$

$$M(n) = M(n-1) + \begin{cases} 0, & \text{if } C(n-1) \neq M(n-1), \\ \xi_n^{(2)}, & \text{if } C(n-1) = M(n-1), \end{cases} \tag{45}$$

for $n \geq 1$.

Let us assume that $M(t) = M([t])$, for non-integer $t > 0$. Clearly, $\{M(t), t \geq 0\}$ is piecewise constant and its trajectories belong to $D[[0,\infty), \mathbb{R}]$.

Litvak and Robert (2012) proved weak convergence

$$\left\{ \frac{1}{\sqrt[4]{n}} M(nt), t \geq 0 \right\} \overset{\mathcal{D}}{\Rightarrow} \left\{ B_1(L_{B_2}(t)), t \geq 0 \right\}, \text{ as } n \to \infty \tag{46}$$

(see Subsection **1.5** for definitions), where $B_1(t)$ and $B_2(t)$ are independent standard Brownian motions on $\mathbb{R}$ and $L_{B_2}(t)$ is the local time process of $B_2(t)$ at 0.

§ 1.12. Cat-and-Mouse model with a general jump distribution of the mouse $(C \to M)$

In this Subsection we introduce our results for Cat-and-Mouse Markov chain with more general distributions of random variables $\xi_n^{(1)}$ and $\xi_n^{(2)}$. We start with the same distribution of $\xi_n^{(1)}$ and generalise distribution of $\xi_n^{(2)}$. Thus, the cat is a simple random walk and the the mouse is a general random walk. Then we also generalise the distribution of $\xi_n^{(1)}$, however we need certain restrictions on the mouse (finite second moments).

1.3.1 We continue to assume that the dynamics of the cat is described by a simple random walk on $\mathbb{Z}$. Let ξ take values ± 1 with equal probabilities. Let $C(0) = 0$, $C(n) = C(n-1) + \xi_n^{(1)}$, where $\xi, \xi_1^{(1)}, \xi_2^{(1)}, \dots$ are independent and identically distributed random variables.

Let $M(0) = 0$, $M(n) = M(n-1) + \xi_n^{(2)} I[C(n-1) = M(n-1)]$ where $\{\xi_n^{(2)}\}_{n=1}^{\infty}$ are independent and identically distributed random variables independent of $\{\xi_n^{(1)}\}_{n=1}^{\infty}$. Assume that

$$\mu = \mathbb{E}\xi_1^{(2)} \text{ is finite.} \tag{47}$$

We do not necessarily need a finite second moment in this case. However we need the distribution of $\xi_1^{(2)} - \mu$ to be in the strict domain of attraction of some stable distribution (see Subsection **1.7** for definitions). If the variance is finite it holds. In the infinite variance case we may ask for the tail-distribution of $\xi_1^{(2)}$ to be regularly varying with index $\alpha \in [1, 2]$. However, for the case $\alpha = 1$ the domain of attraction is not necessarily strict and additional centering might be needed. Thus, we assume that there exist a function $b(c) > 0$, $c \geq 0$, and a random variable $A^{(2)}$ having a stable distribution with index $\alpha \in [1, 2]$ such that

$$\frac{\sum_{k=1}^{n} (\xi_k^{(2)} - \mu)}{b(n)} \Rightarrow A^{(2)}, \text{ as } n \to \infty. \tag{48}$$

Define

$$\tau(0) = 0 \text{ and } \tau(n) = \inf\{m > \tau(n-1) : \ C(m) = M(m)\}, \text{ for } n \geq 1. \tag{49}$$

Given (47), we show that the tail-distribution of $\tau(1)$ is regularly varying with index $1/2$. In the original model, the distance between the cat and the mouse right after a separation equals 2. Now this distance is in form of $\xi_n^{(2)} \pm 1$. Nevertheless, using the theory from Subsection **1.3** we achieve the result. As a consequence of this result, there exists a random variable $D^{(2)}$ having a stable distribution with index $1/2$ such that

$$\frac{\tau(n)}{n^2} \Rightarrow D^{(2)}, \text{ as } n \to \infty. \tag{50}$$

In the proof of Theorem 1.19 we will show that, in fact, there is a joint convergence

$$\left(\frac{\sum_{k=1}^{n} (\xi_k^{(2)} - \mu)}{b(n)}, \ \frac{\tau(n)}{n^2} \right) \Rightarrow (A^{(2)}, D^{(2)}), \ \text{as } n \to \infty, \tag{51}$$

where the random variables on the right-hand side are independent. Further, let $\{(A^{(2)}(t), D^{(2)}(t))\}_{t \geq 0}$ denote a stochastic process with independent increments such that $(A^{(2)}(1), D^{(2)}(1))$ has the same distribution as $(A^{(2)}, D^{(2)})$ or *Lévy process* generated by $(A^{(2)}, D^{(2)})$ (see Subsection **1.7** for further comments). Let $E^{(2)}(s) = \inf\{t \geq 0 : \ D^{(2)}(t) > s\}$.

Theorem 1.19. *Assume that (47) and (48) hold. Then*

- *if $\mu = 0$, we have*

$$\left\{ \frac{M(nt)}{b(\sqrt{n})}, \ t \geq 0 \right\} \overset{\mathcal{D}}{\Rightarrow} \{A^{(2)}(E^{(2)}(t)), t \geq 0\}, \ \text{as } n \to \infty, \tag{52}$$

- *if $\mu \neq 0$, we have*

$$\left\{ \frac{M(nt)}{\sqrt{n}}, \ t \geq 0 \right\} \overset{\mathcal{D}}{\Rightarrow} \{\mu E^{(2)}(t), t \geq 0\}, \ \text{as } n \to \infty. \tag{53}$$

This result illustrate a very simple relation. Up to time n, the number of meeting time-instants (and the jumps of the mouse) is of order $\sqrt{n}$. Thus, the "default scaling" $b(n)$ must be changed to $b(\sqrt{n})$. For the same reason, we get $E^{(2)}(t)$ in the limit.

1.3.2 Assume now that both $\xi_1^{(1)}$ and $\xi_1^{(2)}$ have general distributions on the integer lattice. The main difference for the mouse is that we need to assume finite second moment for $\xi_1^{(2)}$. For further comments on future generalisations of this case see Section **8**. The core of our result is the fact that changing simple random walk to a general random walk does not change the scaling if we assume aperiodicity and finite second moments for the increments.

Theorem 1.20. *Assume that $\mathbb{E}\xi^{(1)} = 0$, $\mathbf{Var}\xi_1^{(1)} < \infty$ and $\xi_1^{(1)}$ has an aperiodic distribution. Assume $\mathbf{Var}\xi_1^{(2)} < \infty$ and, therefore, (48) holds with $b(n) = \sqrt{n\mathbf{Var}\xi_1^{(2)}}$ and a standard normal random variable $A^{(2)}$. Then the statements* (52) *and* (53) *of Theorem **1.19** continue to hold, with $b(\sqrt{n}) = n^{1/4}\sqrt{\mathbf{Var}\xi_1^{(2)}}$ in* (52).

§ 1.13. Dog-and-Cat-and-Mouse model $(D \to C \to M)$

Assume that, as before, ξ takes values ± 1 with equal probabilities $1/2$. Let $\{\xi_n^{(1)}\}_{n=1}^{\infty}$, $\{\xi_n^{(2)}\}_{n=1}^{\infty}$ and $\{\xi_n^{(3)}\}_{n=1}^{\infty}$ be mutually independent sequences of independent copies of ξ. We can define the dynamics of Dog-and-Cat-and-Mouse Markov chain $\{(D(n), C(n), M(n))_n\}_{n=1}^{\infty}$ as follows:

$$D(n) = D(n-1) + \xi_n^{(1)}, \tag{54}$$

$$C(n) = C(n-1) + \begin{cases} 0, & \text{if } D(n-1) \neq C(n-1), \\ \xi_n^{(2)}, & \text{if } D(n-1) = C(n-1), \end{cases} \tag{55}$$

$$M(n) = M(n-1) + \begin{cases} 0, & \text{if } C(n-1) \neq M(n-1), \\ \xi_n^{(3)}, & \text{if } C(n-1) = M(n-1), \end{cases} \tag{56}$$

for $n \geq 1$.

Let $T^{(3)}(0) = 0$ and $T^{(3)}(k) = \min\{n > T^{(3)}(k-1) : \ D(n) = C(n) = M(n)\}$, for $k \geq 1$. We show that the tail-distribution of $T^{(3)}(1)$ is regularly varying with index $1/4$. Thus, there exists a positive random variable $D^{(3)}$ with a stable distribution with index $1/4$ such that

$$\frac{T^{(3)}(k)}{k^4} \Rightarrow D^{(3)}, \text{ as } k \to \infty. \tag{57}$$

Let $\{D^{(3)}(t)\}_{t \geq 0}$ be a Lévy process generated by $D^{(3)}$ and $E^{(3)}(s) = \inf\{t \geq 0 : D^{(3)}(t) > s\}$.

Theorem 1.21. *There exists a positive constant $c > 0$ such that*

$$\left\{ \frac{M(nt)}{cn^{1/8}}, t \geq 0 \right\} \overset{\mathcal{D}}{\Rightarrow} \{B(E^{(3)}(t)), t \geq 0\}, \ \textit{as } n \to \infty, \tag{58}$$

where $B(t)$ is a standard Brownian motion, independent of $E^{(3)}(t)$.

§ 1.14. Linear hierarchical chains $(X_1 \to X_2 \to \ldots \to X_N)$ of length N

In this Subsection we consider a generalisation of the CM Markov chain to the case of N dimensions. Due to the complexity of sample paths for $N > 3$, we did not obtain an analogue of (50) and (57). Instead, we have proved the convergence for every fixed $t > 0$.

Let ξ take values ± 1 with equal probabilities $1/2$. Let $\{\{\xi_n^{(j)}\}_{n=1}^{\infty}\}_{j=1}^{N}$ be mutually independent sequences of independent copies of ξ. Assume $X_1(0) = \ldots = X_N(0) = 0$. Then Markov chain $(X_1(n), \ldots, X_N(n))$ is defined as follows:

$$X_1(n) = X_1(n-1) + \xi_n^{(1)}, \tag{59}$$

$$X_j(n) = X_j(n-1) + \begin{cases} 0, & \text{if } X_{j-1}(n-1) \neq X_j(n-1), \\ \xi_n^{(j)}, & \text{if } X_{j-1}(n-1) = X_j(n-1), \end{cases} \tag{60}$$

for $j \in [2, N]$ and for $n \geq 1$.

Theorem 1.22. *There exists a non-degenerate random variable ζ_N such that, for any fixed $t > 0$,*

$$\frac{X_N([nt])}{n^{2^{-N}}} \Rightarrow t^{2^{-N}} \zeta_N, \text{ as } n \to \infty. \tag{61}$$

Thus, if one agent of the hierarchy has a scaling n^{α}, then the following agent has the scaling $\sqrt{n^{\alpha}}$.

SECTION 2

Linear hierarchical chain of length N

In this section we give a general analysis of trajectories of $(X_1, \ldots, X_N)$ in the case where the increments $\xi_k^{(j)}$ take values ± 1 w.p. $1/2$. In our model if $X_j(n) \neq X_{j-1}(n)$, for $j \in \{2, \ldots, N\}, n \geq 1$, then the j-th coordinate remains unchanged at the time $n+1$. To analyse the dynamics of $\{X_j(n)\}_{n=0}^{\infty}$ it is natural to consider the time instants n when the j-th coordinate changes its value (for such n we have $X_j(n-1) = X_{j-1}(n-1)$) and the corresponding increments $\xi_k^{(j)}$. For $j \in \{1, \ldots, N\}$, let

$$T_j(0) = 0 \text{ and } T_j(k) = \inf\{n > T_j(k-1): \ X_j(n) \neq X_j(n-1)\}, \text{ for } k \geq 1. \quad (62)$$

R.v. $T_j(k)$ is the k-th time-instant when the j-th coordinate changes its value. Since the first coordinate X_1 changes its value every time instant we have $T_1(k) = k$. We consider initial conditions $X_1(0) = \ldots = X_N(0) = 0$. Since the increments $\{\{\xi_k^{(j)}\}_{k=1}^{\infty}\}_{j=1}^{N}$ are non-zero, we have $T_1(1) = \ldots = T_N(1) = 1$. It is important for us to know that, for any $j \in \{1, \ldots, N\}$, the process $\{X_j(T_j(k))\}_{k=0}^{\infty}$ has the same distribution as $\{X_1(n)\}_{n=0}^{\infty}$.

Let us define $\nu_j(n) = \max\{k \geq 0: \ T_j(k) \leq n\}$, the number of time-instants up to time n when X_j changed its value. Then we can rewrite the dynamics of the j-th coordinate as

$$X_j(n) = \sum_{k=1}^{\nu_j(n)} \xi_{T_j(k)}^{(j)}.$$

Our restrictions on the distribution of the increments $\xi_k^{(j)}$, for $k \geq 1$, give us the next important property of our process

Proposition 2.1. *Sequences $\{T_j(k)\}_{k=0}^{\infty}$ and $\{\xi_k^{(j)}\}_{k=1}^{\infty}$ are independent for any $j \in \{1, \ldots, N\}$.*

This property comes from the space-symmetry of the model.

For $j = 1$ the result is trivial, since $\nu_1(n) = n$. We show the result for $j = 2$

and then extend it onto $j > 2$. Define

$$^1\tau(0) = 0 \text{ and } ^1\tau(k) = \inf\{n > {}^1\tau(k-1) : \ X_1(n) = X_2(n)\}, \text{ for } k \geq 1. \tag{63}$$

One can see that in our model $T_2(k) = 1 + {}^1\tau(k-1)$, for $k \geq 1$. In the time interval $[1, {}^1\tau(1)]$ the second coordinate changes its value only at the time $T_2(1) = 1$. Thus, the time ${}^1\tau(1)$ does not depend on $\xi_k^{(2)}$, for $k \geq 2$. Additionally, the trajectory $\{X_1(n)\}_{n=0}^{\infty}$ has the same distribution as $\{-X_1(n)\}_{n=0}^{\infty}$. Thus,

$$\mathbb{P}\{{}^1\tau(1) = n, \ \xi_1^{(2)} = 1\} = \mathbb{P}\{{}^1\tau(1) = n, \ \xi_1^{(2)} = -1\}. \tag{64}$$

As a corollary of the last equation, we get that ${}^1\tau(1)$ has the same distribution as the time that is needed for the simple random walk to hit 0 if it starts from 1. This implies (see Subsection **1.3.3**) that

$$\mathbb{P}\{{}^1\tau(1) > n\} \sim \sqrt{\frac{2}{\pi n}}, \text{ as } n \to \infty. \tag{65}$$

From the symmetry of our model, it follows further that the sequence $\{{}^1\tau(k)\}_{k=0}^{\infty}$, and subsequently the sequences $\{T_2(k)\}_{k=0}^{\infty}$ and $\{\nu_2(n)\}_{n=1}^{\infty}$, do not depend on $\{\xi_k^{(2)}\}_{k=1}^{\infty}$ (and, therefore, on $\{\xi_k^{(j)}\}_{k\geq 1, j\geq 2}$).

For the analysis of $\{T_j(k)\}_{k=0}^{\infty}$, $j > 2$, we need to define an 'embedded version' of ${}^1\tau(k)$. Let

$$^j\tau(0) = 0 \text{ and } ^j\tau(k) = \inf\{m > {}^j\tau(k-1) : \ X_j(T_j(m)) = X_{j+1}(T_j(m))\}, \text{ for } k \geq 1. \tag{66}$$

The process $\{{}^j\tau(k)\}_{k=0}^{\infty}$ counts the number of times that the process X_j changed its value between the time-instants when X_j and X_{j+1} have the same value. Using the same argument as before, we get that the sequence $\{{}^j\tau(k)\}_{k=0}^{\infty}$ does not depend on $\{\xi_k^{(j+1)}\}_{k=1}^{\infty}$.

The j-th coordinate X_j changes its value for the k-th time at the time-instant n if and only if up to time $n-1$ processes X_{j-1} and X_j had the same value exactly $k-1$ times (not including $X_{j-1}(0) = X_j(0) = 0$) and the last time was at the time-instant $n-1$ (which also means that at the time-instant $n-1$ the process X_{j-1} changes its value). This can be rewritten as

$$T_j(k) = n \ \Leftrightarrow \ n - 1 = T_{j-1}({}^{j-1}\tau(k-1)), \text{ for } j \geq 2, \ k \geq 1, \tag{67}$$

and thus $T_j(k) = 1 + T_{j-1}({}^{j-1}\tau(k-1))$. Thus, since sequences $\{T_2(k)\}_{k=0}^{\infty}$ and $\{{}^2\tau(k)\}_{k=0}^{\infty}$ do not depend on $\{\xi_k^{(j)}\}_{k\geq 1, j\geq 3}$, the same holds for $\{T_3(k)\}_{k=0}^{\infty}$. Therefore, using the induction, we get that the sequences $\{T_j(k)\}_{k=0}^{\infty}$ and $\{\xi_k^{(j)}\}_{k=1}^{\infty}$ are independent for any $j \geq 1$.

As a corollary of this result we get

$$X_j(n) = \sum_{k=1}^{\nu_j(n)} \xi_{T_j(k)}^{(j)} \stackrel{d}{=} \sum_{k=1}^{\nu_j(n)} \xi_k^{(j)}. \tag{68}$$

Let ${}^j\eta(n) = \max\{k \geq 0 : \ {}^j\tau(k) \leq n\}$ for $n \geq 0$ and $j \in [1, \ldots, N]$. Since the sequence $\{{}^j\tau(k)\}_{k=0}^{\infty}$ depends only on the sequence $\{\xi_k^{(j)}\}_{k=1}^{\infty}$, we have that $\{{}^j\eta(n)\}_{j=1}^{N-1}$ are i.i.d. r.v.'s. For $n \geq 1$ and $j \in \{1, \ldots, N\}$ we have

$$\begin{aligned}
\nu_j(n) = \max\{k \geq 0 : \ T_j(k) \leq n\} &= \max\{k \geq 1 : \ 1 + T_{j-1}({}^{j-1}\tau(k-1)) \leq n\} \\
&= 1 + \max\{k \geq 0 : \ T_{j-1}({}^{j-1}\tau(k)) \leq n-1\} \\
&= 1 + \max\{k \geq 0 : \ {}^{j-1}\tau(k) \leq \nu_{j-1}(n-1)\} \\
&= 1 + {}^{j-1}\eta(\nu_{j-1}(n-1)).
\end{aligned} \tag{69}$$

For $n < N-1$ we can iterate the process and get

$$\begin{aligned}
\nu_N(n) &= 1 + {}^{N-1}\eta(1 + {}^{N-2}\eta(\ldots(1 + {}^{N-n}\eta(0))\ldots)) \\
&= 1 + {}^{N-1}\eta(1 + {}^{N-2}\eta(\ldots(1 + {}^{N-n+1}\eta(1))\ldots)) \\
&\stackrel{d}{=} 1 + {}^{n-1}\eta(1 + {}^{n-2}\eta(\ldots(1 + {}^{1}\eta(1))\ldots)) \\
&= \nu_n(n).
\end{aligned} \tag{70}$$

For $n \geq N-1$ we have

$$\nu_N(n) = 1 + {}^{N-1}\eta(1 + {}^{N-2}\eta(\ldots + {}^{1}\eta(n-N+1))). \tag{71}$$

We want to construct a process with the same distribution as $\{\nu_N(n)\}_{n=0}^{\infty}$ in a form of $\nu_{N-1}(\varphi(n))$, where process$\{\varphi(n)\}_{n=0}^{\infty}$ is independent of everything else. Define process $\{\eta(n)\}_{n=0}^{\infty} \stackrel{d}{=} \{{}^{N-1}\eta(n)\}_{n=0}^{\infty}$, which is independent of everything else. Then, for $n \geq N-1$, we have

$$\nu_N(n) \stackrel{d}{=} 1 + {}^{N-2}\eta(1 + {}^{N-3}\eta(\ldots + \eta(n-N+1))). \tag{72}$$

Using the same formula for $\nu_{N-1}(m)$ with such m that $m - (N-1) + 1 = 1 + {}^{N-1}\eta(n-N+1)$, we get

$$\nu_N(n) \stackrel{d}{=} \nu_{N-1}(N - 1 + \eta(n-N+1)), \text{ for } n \geq N-1. \tag{73}$$

Then, for $n \geq N$ we have $X_N(n) \overset{d}{=} X_{N-1}(N-1+\eta(n-N+1))$. There exists a non-degenerate r.v. ζ (see Subsection**1.3.3**) such that

$$\mathbb{P}\{^{N-1}\tau(1) > n\}\eta(n) \Rightarrow \zeta, \text{as } n \to \infty. \tag{74}$$

Therefore, using (65) we get

$$\frac{j-1+\eta(n-j+1)}{\sqrt{n}} = \frac{j-1+\eta(n-j+1)}{\sqrt{n-j+1}} \frac{n-j+1}{\sqrt{n}} \Rightarrow \sqrt{\frac{\pi}{2}}\zeta, \tag{75}$$

as $n \to \infty$, for $j \geq 1$. We now present a known result that we utilise to prove Theorem **1.22**.

Proposition 2.2. *(Dobrushin (1955), (v)) Let $Y(t)$ and τ_n be independent sequences of r.v.'s such that*

$$\frac{Y(t)}{bt^{\beta}} \Rightarrow Y, \ \text{as } t \to \infty, \ \text{and} \ \frac{\tau_n}{dn^{\delta}} \Rightarrow \tau, \ \text{as } n \to \infty. \tag{76}$$

Then for independent Y and τ we have

$$\frac{Y(\tau_n)}{bd^{\beta}n^{\beta\delta}} \Rightarrow Y\tau^{\beta}, \ \text{as } n \to \infty. \tag{77}$$

Indeed, by the Central Limit Theorem $X_1(n)/\sqrt{n}$ weakly converges to a normally distributed r.v. ψ (we assume that ψ and ζ are independent). Together with (75) and independence of $X_1(n)$ and $\eta(n)$, this insures that condition (76) holds with $Y(t) = X_1([t])$, $\tau_n = 1+\eta(n-1)$ and $\beta = \delta = 1/2$. By the Proposition **2.2**, we get

$$\frac{X_2(n)}{n^{1/4}} \overset{d}{=} \frac{X_1(1+\eta(n-1))}{n^{1/4}} \Rightarrow \psi\sqrt{\sqrt{\frac{\pi}{2}}\zeta}, \text{ as } n \to \infty. \tag{78}$$

Let $\{\zeta_j\}_{j=2}^N$ be independent copies of ζ which are independent of ψ. Next we use the induction argument. For some $j \geq 1$ we have that condition (76) holds with $Y(t) = X_j([t])$, $\tau_n = j-1+\eta(n-j+1)$, $\beta = 2^{-j}$ and $\delta = 2^{-1}$. By the Proposition **2.2**, we get

$$\frac{X_{j+1}(n)}{n^{2^{-(j+1)}}} \overset{d}{=} \frac{X_j(j-1+\eta(n-j+1))}{n^{2^{-(j+1)}}} \Rightarrow \psi\prod_{i=2}^{j+1}\sqrt{\sqrt{\frac{\pi}{2}}\zeta_i}, \text{ as } n \to \infty. \tag{79}$$

This concludes the proof Theorem **1.22**.

SECTION 3

The Cat-and-Mouse model

In this section we consider the two-dimensional version of the model and prove Theorem **1.19** and Theorem **1.20** which establish the asymptotic behaviour of the second coordinate. In Section **2** we presented the basic analysis of the trajectories. Using its notations we get

$$M(0) = 0, \ M(n) = \sum_{k=1}^{1+{}^1\eta(n-1)} \xi^{(2)}_{T_2(k)}, \text{ for } n \geq 1. \tag{80}$$

Since we are interested only in the weak asymptotic behaviour, we can properly define another stochastic process having the same distribution as $\{M(n)\}_{n=0}^{\infty}$ and prove our result for a new process.

For $i = 1, 2$, let $S_0^{(i)} = 0$ and $S_n^{(i)} = \sum_{k=1}^{n} \xi_k^{(i)}$, for $n \geq 1$. Let

$$\tau(0) = 0 \text{ and } \tau(n) = \inf\{m > \tau(n-1) : \ S_m^{(1)} = S_n^{(2)}\}, \text{ for } n \geq 1. \tag{81}$$

Since $\{\xi_k^{(1)}\}_{k=1}^{\infty}$ and $\{\xi_k^{(2)}\}_{k=1}^{\infty}$ are independent sequences of i.i.d. r.v.'s, we have that $\tau(n) - \tau(n-1) \stackrel{d}{=} \tau(1)$, for $n \geq 1$.

Let $\eta(t) = \max\{k \geq 0 : \ \tau(k) \leq t\}$ for $t \geq 0$. Define a continuous-time process $M'(t)$ by

$$M'(t) = 0 \text{ for } t \in [0, 1) \text{ and } M'(t) = S^{(2)}_{\eta(t-1)+1} = \sum_{k=1}^{\eta(t-1)+1} \xi_k^{(2)}, \text{ for } t \geq 1. \tag{82}$$

It is straightforward to check that $\{M'(n), \ n \geq 0\} \stackrel{d}{=} \{M(n), \ n \geq 0\}$. In the rest of the section we will omit the dash and simply write $M(t)$. The process $\{\widehat{M}(t)\}_{t\geq 0} = \{M'(t+1)\}_{t\geq 0}$ is a so-called *oracle continuous-time random walk* (see, e.g., Jurlewicz *et al.* (2010)). We need the following proposition (see Subsection **1.5** for more details).

Proposition 3.1. *We have*

$$d_{\mathcal{J}_1,\infty}\left(\left\{\frac{M(ct)}{b(\sqrt{c})}, t \geq 0\right\}, \left\{\frac{M(ct+1)}{b(\sqrt{c})}, t \geq 0\right\}\right) \stackrel{a.s.}{\to} 0, \ \textit{as } c \to \infty. \tag{83}$$

First, consider the case $\mathbb{E}\xi_1^{(2)} = 0$. We want to show that

$$\left(\frac{S_n^{(2)}}{b(n)}, \frac{\tau(n)}{n^2}\right) \Rightarrow (A^{(2)}, D^{(2)}), \text{ as } n \to \infty. \tag{84}$$

Given that, we will show that the first part of Theorem **1.19** follows from the next proposition.

Proposition 3.2. *(Theorem 3.1, Jurlewicz et al. (2010)) Assume (84) holds. Then*

$$\left\{\frac{\widehat{M}(ct)}{b(\sqrt{c})}, t \geq 0\right\} = \left\{\frac{M(ct+1)}{b(\sqrt{c})}, t \geq 0\right\} \overset{\mathcal{D}}{\Rightarrow} \left\{A^{(2)}(E^{(2)}(t)), t \geq 0\right\}, \text{ as } c \to \infty. \tag{85}$$

We will now show that relation (84) holds and that r.v.'s $A^{(2)}$ and $D^{(2)}$ are independent, which means

$$\begin{aligned}\mathbb{E}\exp\left(i\left(\lambda_1\frac{S_n^{(2)}}{b(n)} + \lambda_2\frac{\tau(n)}{n^2}\right)\right) &= \mathbb{E}\exp\left(i\left(\lambda_1\frac{\xi_1^{(2)}}{b(n)} + \lambda_2\frac{\tau(1)}{n^2}\right)\right)^n \\ &= \left(1 + \frac{f_1(\lambda_1) + f_2(\lambda_2)}{n} + o\left(\frac{1}{n}\right)\right)^n, \end{aligned} \tag{86}$$

as $n \to \infty$, for some function f_1 and f_2. Indeed, convergence of characteristic functions is equivalent to weak convergence of r.v.'s and for independence of r.v.'s it is sufficient to verify that the characteristic function of the sum is equal to the product of respective characteristic functions. Since the right-hand side of (86) converges to $\exp(f_1(\lambda_1))\exp(f_2(\lambda_2))$, it will prove the convergence and the independence of the limits $A^{(2)}$ and $D^{(2)}$.

3.1 We start with the case of Theorem **1.19**. From (48) we have a weak convergence to a random variable. Again, this is equivalent to convergence of characteristic functions. Thus, (48) implies

$$\mathbb{E}\exp\left(i\lambda_1\frac{\sum_{k=1}^n \xi_k^{(2)}}{b(n)}\right) = \left[\mathbb{E}\exp\left(i\lambda_1\frac{\xi_1^{(2)}}{b(n)}\right)\right]^n \to \mathbb{E}\exp\left(i\lambda_1 A^{(2)}\right), \text{ as } n \to \infty. \tag{87}$$

Additionally, if $B^n(n) \to z$, as $n \to \infty$, then $n \log B(n) \to \log z$, which leads to $\log B(n) \sim n^{-1}\log z$. Finally, such relation leads to $B(n) \sim 1 + n^{-1}\log z$, as $n \to \infty$. Thus, we have the following

$$\mathbb{E}\exp\left(i\lambda_1\frac{\xi_1^{(2)}}{b(n)}\right) \sim 1 + \frac{l_1(\lambda_1)}{n}, \text{as } n \to \infty, \tag{88}$$

where $l_1(\lambda) = \log \mathbb{E} \exp(i\lambda A^{(2)})$, the logarithmic characteristic function of $A^{(2)}$.

Let $\{\tau_k^{(1)}\}_{k=1}^{\infty}$ be independent copies of τ, the time needed for the simple random walk to hit 0 if it starts from 1, independent of $\{\xi_n^{(2)}\}_{n=1}^{\infty}$. Then we have the following relation for $\tau(1)$:

$$\tau(1) \stackrel{d}{=} I[\xi_n^{(2)} \neq 0] \sum_{k=1}^{|\xi_n^{(2)}|} \tau_k^{(1)} + I[\xi_n^{(2)} = 0](1 + \tau_1^{(1)}). \tag{89}$$

Since $\mathbb{P}\{\tau > n\} \sim \sqrt{2/(\pi n)}$, as $n \to \infty$, a combination of Propositions **1.8** and **1.11** from Subsection **1.3** gives us

$$\mathbb{P}\{\tau(1) > n\} \sim (\mathbb{E}|\xi_1^{(2)}| + \mathbb{P}\{\xi_1^{(2)} = 0\})\mathbb{P}\{\tau > n\} \tag{90}$$

and there exists a positive stable law $D^{(2)}$ such that

$$\frac{\tau(n)}{n^2} \Rightarrow D^{(2)}, \text{ as } n \to \infty \tag{91}$$

(see Subsection **1.7**).

Using the same argument as for (88) we get

$$\mathbb{E} \exp\left(i\lambda_2 \frac{\tau}{n^2}\right) \sim 1 + \frac{l_2(\lambda_2)}{n}, \text{as } n \to \infty, \tag{92}$$

where $\lambda_2(\lambda) = \log \mathbb{E} \exp(i\lambda D^{(2)}/(\mathbb{E}|\xi_1^{(2)}| + \mathbb{P}\{\xi_1^{(2)} = 0\}))$, the logarithmic characteristic function of $D^{(2)}/(\mathbb{E}|\xi_1^{(2)}| + \mathbb{P}\{\xi_1^{(2)} = 0\})$. Then we have

$$\begin{aligned}
&\mathbb{E} \exp\left(i\left[\lambda_1 \frac{\xi_1^{(2)}}{b(n)} + \lambda_2 \frac{\tau(1)}{n^2}\right]\right) \\
&\qquad = \sum_{-\infty}^{\infty} \exp\left(i\lambda_1 \frac{k}{b(n)}\right) \mathbb{P}\{\xi_1^{(2)} = k\} \mathbb{E} \exp\left(i\lambda_2 \frac{\tau(1)}{n^2} |\xi_1^{(2)} = k\right) \\
&\qquad\qquad = \mathbb{P}\{\xi_1^{(2)} = 0\} \mathbb{E} \exp\left(i\lambda_2 \frac{1+\tau}{n^2}\right) + \\
&\qquad\qquad + \sum_{k \neq 0} \exp\left(i\lambda_1 \frac{k}{b(n)}\right) \mathbb{P}\{\xi_1^{(2)} = k\} \left(\mathbb{E} \exp\left(i\lambda_2 \frac{\tau}{n^2}\right)\right)^{|k|}.
\end{aligned} \tag{93}$$

Using (92), for any $m > 0$

$$\begin{aligned}
\left(\mathbb{E} \exp\left(i\lambda_2 \frac{\tau}{n^2}\right)\right)^m &= \left(1 + \frac{l_2(\lambda_2)}{n} + o\left(\frac{1}{n}\right)\right)^m \\
= \exp\left(m \ln(1 + \frac{l_2(\lambda_2) + o(1)}{n})\right) &= \exp\left(\frac{m l_2(\lambda_2)}{n}(1 + o(1))\right) \\
= 1 + \frac{m l_2(\lambda_2)(1 + o(1))}{n} &+ \frac{1}{n^2} \sum_{j=2}^{\infty} \frac{(m l_2(\lambda_2)(1 + o(1)))^j}{n^{j-2} j!},
\end{aligned} \tag{94}$$

as $n \to \infty$. Now we use the fact that if $\sum_{-\infty}^{\infty} A_n = \sum_{-\infty}^{\infty} B_n + \sum_{-\infty}^{\infty} C_n$ and if series $\sum_{-\infty}^{\infty} A_n$ and $\sum_{-\infty}^{\infty} B_n$ converge, then $\sum_{-\infty}^{\infty} C_n$ converges too. We have

$$\sum_{k\neq 0} \exp\left(i\lambda_1 \frac{k}{b(n)}\right) \mathbb{P}\{\xi_1^{(2)} = k\} \left(\mathbb{E}\exp\left(i\lambda_2 \frac{\tau}{n^2}\right)\right)^{|k|}$$
$$= \left(\mathbb{E}\exp\left(i\lambda_1 \frac{\xi_1^{(2)}}{b(n)}\right) - \mathbb{P}\{\xi_1^{(2)} = 0\}\right)$$
$$+ \frac{l_2(\lambda_2)}{n} \sum_{k\neq 0} |k| \exp\left(i\lambda_1 \frac{k}{b(n)}\right) \mathbb{P}\{\xi_1^{(2)} = k\} + o\left(\frac{1}{n}\right)$$
$$= \left(\mathbb{E}\exp\left(i\lambda_1 \frac{\xi_1^{(2)}}{b(n)}\right) - \mathbb{P}\{\xi_1^{(2)} = 0\}\right) + \mathbb{E}|\xi_1^{(2)}| \frac{l_2(\lambda_2)}{n} + o\left(\frac{1}{n}\right), \quad (95)$$

as $n \to \infty$. Using (88) and (92), we have

$$\mathbb{E}\exp\left(i\left[\lambda_1 \frac{\xi_1^{(2)}}{b(n)} + \lambda_2 \frac{\tau(1)}{n^2}\right]\right)$$
$$= 1 + \frac{l_1(\lambda_1)}{n} + (\mathbb{E}|\xi_1^{(2)}| + \mathbb{P}\{\xi_1^{(2)} = 0\}) \frac{l_2(\lambda_2)}{n} + o\left(\frac{1}{n}\right), \quad (96)$$

as $n \to \infty$. We have proved that equation (86) holds with $f_1(\lambda_1) = l_1(\lambda_1)$ and $f_2(\lambda_2) = (\mathbb{E}|\xi_1^{(2)}| + \mathbb{P}\{\xi_1^{(2)} = 0\}) l_2(\lambda_2)$. Therefore, equation (84) holds and we can use Propositions **3.1** and **3.2** to prove the first part of Theorem **1.19**.

Turn now to the second part and assume $\mathbb{E}\xi^{(2)} = \mu \neq 0$. Then the above arguments are applicable to $\sum_{k=1}^{\eta(t)+1} (\xi_k^{(2)} - \mu)$. Thus, we have shown that the process $\left(\left(\sum_{k=1}^{\eta(nt)+1} (\xi_k^{(2)} - \mu)\right) / b(\sqrt{n}),\ t \geq 0\right)$ weakly converges to the limiting one (see Subsection **1.5** for corresponding definitions). Since $\mu < \infty$, we have $b(n) = o(n)$, as $n \to \infty$, and therefore the process

$$\left(\frac{\sum_{k=1}^{\eta(nt)+1} (\xi_k^{(2)} - \mu)}{\sqrt{n}},\ t \geq 0\right) = \left(\frac{\sum_{k=1}^{\eta(nt)+1} (\xi_k^{(2)} - \mu)}{b(\sqrt{n})} \frac{b(\sqrt{n})}{\sqrt{n}},\ t \geq 0\right) \quad (97)$$

converges to the zero-valued process.

Thus, it follows from the representation

$$\frac{\widehat{M}(nt)}{\sqrt{n}} = \frac{\sum_{k=1}^{\eta(nt)+1} (\xi_k^{(2)} - \mu)}{\sqrt{n}} + \frac{\mu(\eta(nt) + 1)}{\sqrt{n}} \quad (98)$$

and from the Corollary of Theorem 3.2 from Meerschaert and Scheffler (2004) (see also Appendix **A.2**) that

$$\left\{\frac{\widehat{M}(nt)}{\sqrt{n}},\ t \geq 0\right\} \overset{\mathcal{D}}{\Rightarrow} \{\mu E^{(2)}(t),\ t \geq 0\} \text{ as } n \to \infty. \quad (99)$$

3.2 Turn now to the second case. Under the assumption of the finiteness of second moments we can expand our result to the case where both $\xi^{(1)}$ and $\xi^{(2)}$ have general distributions. Assume now that $\{S_n^{(1)}\}_{n=0}^{\infty} = \{\sum_{k=1}^{n} \xi_k^{(1)}\}_{n=0}^{\infty}$ is an aperiodic random walk with zero-mean and finite-variance-σ_1^2 increments. A theory of general random walks and their hitting times is quite developed. Nevertheless, it was challenging for us to find results uniformal in terms of the hitting point. From Section **3.3** of Uchiyama (2011a), we have that, uniformly in x,

$$\mathbb{E}\left[\exp\left(it\tau(1)\right) \mid \xi_1^{(2)} = x\right] = 1 - (a^*(x) + e_x(t))(\sigma_1\sqrt{-2it} + o(\sqrt{|t|})), \text{ as } t \to 0, \tag{100}$$

where

$$a^*(x) = 1 + \sum_{n=1}^{\infty} \left(\mathbb{P}\{S_n^{(1)} = 0\} - \mathbb{P}\{S_n^{(1)} = -x\}\right), \tag{101}$$

$$e_x(t) = c_x(t) + is_x(t), \tag{102}$$

$$|c_x(t)| = O\left(x^2\sqrt{|t|}\right), \text{ as } t \to 0, \tag{103}$$

$$s_0(t) = 0 \text{ and } \frac{s_x(t)}{x} = o(1), \text{ as } t \to 0, \text{ uniformly in } x \neq 0. \tag{104}$$

Following the similar steps from the previous part we take $t = \lambda_2/n^2$ and, eventually, let n become large. A very important relation here is (103). When we take characteristic function $\mathbb{E}\exp\left(i\left[\lambda_1\frac{\xi_1^{(2)}}{\sigma_2\sqrt{n}} + \lambda_2\frac{\tau(1)}{n^2}\right]\right)$ and start to separate it into different summands the relation (103) leads to a summand

$$\sum_{x\in\mathbb{Z}} O\left(\frac{x^2}{n^2}\right)\mathbb{P}\{\xi_1^{(2)} = x\}, \text{ as } n \to \infty, \tag{105}$$

and this is the main reason why we need to assume that $\xi_1^{(2)}$ has finite second moment.

Assume now that $\mathbb{E}\xi_1^{(2)} = 0$ and $\sigma_2 = \mathbf{Var}\xi_1^{(2)} < \infty$. We have (see, e.g., Proposition 7.2 from Uchiyama (2011b))

$$\sigma_1^2(a^*(x) - I(x=0)) \sim |x|, \quad \text{as } |x| \to \infty. \tag{106}$$

As a consequence we get $\mathbb{E}a^*(\xi_1^{(2)}) < \infty$. Let $p^{(2)}(x) = \mathbb{P}\{\xi_1^{(2)} = x\}$. Then total probability formula gives us

$$\mathbb{E}\exp\left(i\left[\lambda_1\frac{\xi_1^{(2)}}{\sigma_2\sqrt{n}}+\lambda_2\frac{\tau(1)}{n^2}\right]\right)$$
$$=\sum_{x\in\mathbb{Z}}\exp\left(i\lambda_1\left[\frac{x}{\sigma_2\sqrt{n}}\right]\right)\mathbb{E}\left[\exp\left(i\left[\lambda_2\frac{\tau(1)}{n^2}\right]\right)\mid\xi_1^{(2)}=x\right]p^{(2)}(x). \quad (107)$$

Now we use (100)-(104) to get

$$\mathbb{E}\exp\left(i\left[\lambda_1\frac{\xi_1^{(2)}}{\sigma_2\sqrt{n}}+\lambda_2\frac{\tau(1)}{n^2}\right]\right)=\mathbb{E}\left[\exp\left(i\lambda_1\left[\frac{\xi_1^{(2)}}{\sigma_2\sqrt{n}}\right]\right)\right]-$$
$$-\frac{\sigma_1\sqrt{-2i\lambda_2}}{n}\mathbb{E}\left[a^*(\xi_1^{(2)})\exp\left(i\lambda_1\left[\frac{\xi_1^{(2)}}{\sigma_2\sqrt{n}}\right]\right)\right]+$$
$$+O\left(\frac{1}{n^2}\mathbb{E}\left[\left(\xi_1^{(2)}\right)^2\exp\left(i\lambda_1\left[\frac{\xi_1^{(2)}}{\sigma_2\sqrt{n}}\right]\right)\right]\right)+$$
$$+o\left(\frac{1}{n}\mathbb{E}\left[\xi_1^{(2)}\exp\left(i\lambda_1\left[\frac{\xi_1^{(2)}}{\sigma_2\sqrt{n}}\right]\right)\right]\right)+o\left(\frac{1}{n}\right), \quad (108)$$

as $n\to\infty$. Next, we relation (106) and the Taylor expansion for the exponent to get

$$\mathbb{E}\left[a^*(\xi_1^{(2)})\exp\left(i\lambda_1\left[\frac{\xi_1^{(2)}}{\sigma_2\sqrt{n}}\right]\right)\right]=\mathbb{E}\left[a^*(\xi_1^{(2)})\right]+o(1),\ \text{as } n\to\infty. \quad (109)$$

Since $\mathbb{E}\xi_1^{(2)}=0$ and $\mathbf{Var}\xi_1^{(2)}<\infty$, the Central Limit Theorem holds. Thus, we have the analogue of (88) with l_1 being a logarithmic characteristic function of a random variable with standard normal distribution. Finally, we get

$$\mathbb{E}\exp\left(i\left[\lambda_1\frac{\xi_1^{(2)}}{\sigma_2\sqrt{n}}+\lambda_2\frac{\tau(1)}{n^2}\right]\right)=1+\frac{l_1(\lambda_1)}{n}-\frac{\sigma_1\sqrt{-2i\lambda_2}}{n}\mathbb{E}\left[a^*(\xi_1^{(2)})\right]+o\left(\frac{1}{n}\right). \quad (110)$$

Thus, we proved equation (86) for this case and the rest of the proof follows the same argument as in the previous case.

SECTION 4

The Dog-and-Cat-and-Mouse model

In this Section we study the structural properties of the three-dimensional version of our model where the increments $\xi_k^{(j)}$ take values ± 1 w.p. $1/2$. We use this analysis to prove Theorem **1.21**. Using the notations from Section **2** we get

$$M(0) = 0,\ M(n) = \sum_{k=1}^{\nu_3(n)} \xi^{(3)}_{T_3(k)},\ \text{for } n \geq 1, \tag{111}$$

$$\nu_3(n) = 1 + {}^2\eta(1 + {}^1\eta(n-2)),\ \text{for } n \geq 2. \tag{112}$$

We construct a simpler process $\widetilde{M}$ and use Theorem 5.1 from Kasahara (1984) to obtain its scaling properties. Then we show that M and $\widetilde{M}$ have the same limiting behaviour.

§ 4.1. Comments on the CM model

Here we revisit the 'standard' CM model and highlight a number of properties that are of use in the analysis of the DCM model.

We assume that $C(0) = M(0) = 0$. Let $V_n = |C(n) - M(n)|$, for $n \geq 0$. Then we can write $M(n+1) = M(n) + \xi^{(2)}_{n+1} I[V_n = 0]$, for $n \geq 1$. Note that $V_{n+1} = |C(n+1) - M(n+1)| = |C(n) - M(n) + \xi^{(1)}_{n+1} - \xi^{(2)}_{n+1} I[V_n = 0]|$. We can further observe that

$$\text{if } V_n = 0, \text{ then } \ V_{n+1} = |\xi^{(1)}_{n+1} - \xi^{(2)}_{n+1}| \stackrel{d}{=} 1 + \xi^{(1)}_{n+1}, \tag{113}$$

$$\text{if } V_n \neq 0, \text{ then } \ V_{n+1} = |C(n) - M(n) + \xi^{(1)}_{n+1}| \stackrel{d}{=} V_n + \xi^{(1)}_{n+1}. \tag{114}$$

Thus, V_n forms a Markov chain. Let $p_i(j) = \mathbb{P}\{V_{n+1} = j | V_n = i\}$, for $i, j \geq 0$. Let

$$\widehat{T}^{(2)}(0) = 0 \text{ and } \widehat{T}^{(2)}(k) = \min\{n > \widehat{T}^{(2)}(k-1) :\ V_n \in \{0,1\}\}. \tag{115}$$

Since $p_0(j) = p_1(j)$ for any j, we have that random variables $\{\widehat{T}^{(2)}(k) - \widehat{T}^{(2)}(k-1)\}_{k=1}^{\infty}$ are independent and identically distributed and random variable $(\widehat{T}^{(2)}(k) - \widehat{T}^{(2)}(k-1))$ does not depend on $V_{\widehat{T}^{(2)}(k-1)}$, for $k \geq 1$. Due to the Markov property, we have

$$V_{\widehat{T}^{(2)}(k)+1} \stackrel{d}{=} 1 + \xi_1^{(1)} = \begin{cases} 0, & \text{w.p. } \frac{1}{2}, \\ 2, & \text{w.p. } \frac{1}{2}. \end{cases} \tag{116}$$

Thus, after each time-instant $\widehat{T}^{(2)}(k)$, the cat and the mouse jump with equal probabilities either to the same point or to two different points distant by 2. In the latter case, $V_{\widehat{T}^{(2)}(k+1)} = 1$, since the cat's jumps are 1 or -1. For the cat, let $\tau_m^{(1)} = \min\{n : \sum_{k=1}^{n} \xi_k^{(1)} = m\}$ denote the hitting time of the state m. Then

$$\widehat{T}^{(2)}(1) \stackrel{d}{=} 1 + \begin{cases} 0, & \text{w.p. } \frac{1}{2}, \\ \tau_1^{(1)}, & \text{w.p. } \frac{1}{2}. \end{cases} \tag{117}$$

The tail asymptotics for $\tau_1^{(1)}$ are known: $\mathbb{P}\{\tau_1^{(1)} > n\} \sim \sqrt{2/(\pi n)}$, as $n \to \infty$ (see Subsection **1.3.3**). Since $\tau_1^{(1)}$ has a distribution with a regularly varying tail, for any $m = 2, 3, \ldots$ we have $\mathbb{P}\{\tau_m^{(1)} > n\} \sim m\mathbb{P}\{\tau_1^{(1)} > n\} \sim \sqrt{2m^2/(\pi n)}$ as $n \to \infty$.

§ 4.2. Auxiliary continuous-time random walk in the DCM model

We assume that $D(0) = C(0) = M(0) = 0$. Let

$$T^{(3)}(0) = 0 \text{ and } T^{(3)}(k) = \min(n > T^{(3)}(k-1) : D(n) = C(n) = M(n)), \tag{118}$$

for $k \geq 1$. Let

$$\{J_k^{(3)}\}_{k=1}^{\infty} = \{T^{(3)}(k) - T^{(3)}(k-1)\}_{k=1}^{\infty} \tag{119}$$

and

$$\{Y_k^{(3)}\}_{k=1}^{\infty} = \{M(T^{(3)}(k)) - M(T^{(3)}(k-1))\}_{k=1}^{\infty}. \tag{120}$$

Denote $\eta(t) = \max\{k : \ T^{(3)}(k) \leq t\}$, for $t \geq 0$. Let $S_0 = 0$ and $S_n = \sum_{k=1}^{n} Y_k$. We define process $\widetilde{M}(t)$ by

$$\widetilde{M}(t) = S_{\eta(t)} = \sum_{k=1}^{\eta(t)} Y_k^{(3)}, \text{ for } t \geq 0. \tag{121}$$

The process $\widetilde{M}(t)$ is a so-called *coupled continuous-time random walk* (see Becker-Kern *et al.* (2004))

In the next two Subsections we analyse the distributions of $Y_n^{(3)}$ and $J_n^{(3)}$.

§ 4.3. Distribution of random variable $J_1^{(3)}$

In this Subsection we find the tail asymptotics of the time between the meeting time-instants of all three agents. In order to do so, we partially use the analysis from Subsection **4.1**. The main idea is to introduce an auxiliary state which shows an overall structure of the process.

Let $V_n = (V_{n1}, V_{n2}) = (|D(n) - C(n)|, |C(n) - M(n)|)$. Then we can write

$$\begin{aligned}&(D(n+1), C(n+1), M(n+1))\\&\qquad = (D(n) + \xi_{n+1}^{(1)}, C(n) + \xi_{n+1}^{(2)} I[V_{n1} = 0], M(n) + \xi_{n+1}^{(3)} I[V_{n2} = 0]). \quad (122)\end{aligned}$$

Note further that

$$\text{if } V_{n1} = V_{n2} = 0, \text{ then } \ V_{n+1} = (|\xi_{n+1}^{(1)} - \xi_{n+1}^{(2)}|, |\xi_{n+1}^{(2)} - \xi_{n+1}^{(3)}|) \stackrel{d}{=} (1 + \xi_{n+1}^{(1)}, 1 + \xi_{n+1}^{(2)}), \quad (123)$$

$$\text{if } V_{n1} = 0 \text{ and } V_{n2} \neq 0, \text{ then } \ V_{n+1} \stackrel{d}{=} (1 + \xi_{n+1}^{(1)}, V_{n2} + \xi_{n+1}^{(2)}), \quad (124)$$

$$\text{if } V_{n1} \neq 0 \text{ and } V_{n2} = 0, \text{ then } \ V_{n+1} \stackrel{d}{=} (V_{n1} + \xi_{n+1}^{(1)}, 1), \quad (125)$$

$$\text{if } V_{n1} \neq 0 \text{ and } V_{n2} \neq 0, \text{ then } \ V_{n+1} \stackrel{d}{=} (V_{n1} + \xi_{n+1}^{(1)}, V_{n2}). \quad (126)$$

Thus, V_n is Markov chain. Let $p_{ij}(m,l) = \mathbb{P}\{V_{n+1} = (m,l) | V_n = (i,j)\}$, for $i, j, m, l \geq 0$. Let

$$\widehat{T}^{(3)}(0) = 0 \text{ and } \widehat{T}^{(3)}(k) = \min\{n > \widehat{T}^{(3)}(k-1) : \ V_n \in \{(0,0), (0,1)\}\}. \quad (127)$$

Since $p_{00}(m,l) = p_{01}(m,l)$ for any m, l, we have that random variables $\{\widehat{T}^{(3)}(k) - \widehat{T}^{(3)}(k-1)\}_{k=1}^{\infty}$ are independent and identically distributed and random variable $(\widehat{T}^{(3)}(k) - \widehat{T}^{(3)}(k-1))$ does not depend on $V_{\widehat{T}^{(3)}(k-1)}$, for $k \geq 1$.

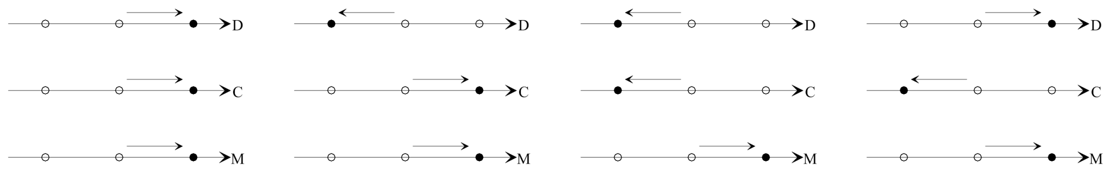
In other words, the auxiliary states are $D(n) = C(n) = M(n) \pm 1$. To find the resulting asymptotics we need the asymptotics of $\widehat{T}^{(3)}(1)$ and the relation between time-instants $T^{(3)}(1)$ and $\{\widehat{T}^{(3)}(k)\}_{k=1}^{\infty}$.

Lemma 4.1. *Let $V_0 \in \{(0,0), (0,1)\}$. Then there exists $c > 0$ such that $\mathbb{P}\{\widehat{T}^{(3)}(1) > n\} \sim c/n^{1/4}$, as $n \to \infty$. Further, $\widehat{T}^{(3)}(1) = 1$ iff $V_{\widehat{T}^{(3)}(1)} = (0,0)$.*

Proof. Let $V_0 = (0,0)$. It is apparent from equation (123) that

$$\mathbb{P}\{V_1 = (0,0)\} = \mathbb{P}\{V_1 = (2,0)\} = \mathbb{P}\{V_1 = (0,2)\} = \mathbb{P}\{V_1 = (2,2)\} = \frac{1}{4}. \quad (128)$$

Since $p_{00}(k,l) = p_{01}(k,l)$, random variable V_1 has the same distribution given $V_0 = (0,1)$.



(a) $V_1 = (0,0)$. (b) $V_1 = (2,0)$. (c) $V_1 = (0,2)$. (d) $V_1 = (2,2)$.

Figure 1: The positioning after the first jump.

Let $V_1 = (0,2)$ (Figure 1c). From equations (124) and (126) we know that $|V_{(k+1)2} - V_{k2}| \in \{0,1\}$, given $V_{k2} \neq 0$, therefore V_{k2} arrives at 1, before hitting 0 and $V_{\widehat{T}^{(3)}(1)} = (0,1)$. Let $\tau, \tau_1, \tau_2, \ldots$ be independent copies of $\tau_1^{(1)}$. Then $\widehat{T}^{(3)}(1)$ has the same distribution as $\sum_{k=1}^{\tau} \tau_k$ and we have that $\mathbb{P}\{\sum_{k=1}^{\tau} \tau_k > n\} \sim c/n^{1/4}$, as $n \to \infty$ (see Subsection **1.3** and Appendix **A.1**).

Let $V_1 = (2,2)$ (Figure 1d). From equation (126), V_{k2} remains at 2 (the cat and the mouse do not move) until V_{k1} reaches 0. This happens after a time which has the same distribution as $\tau_2^{(1)} = \min\{n > 0 : \sum_{k=1}^{n} \xi_k^{(1)} = 2\}$. Thus, we travel from $(2,2)$ to $(0,2)$ while never hitting $(0,0)$. We also know that the tail distribution of the travel time is $\mathbb{P}\{\tau_2^{(1)} > n\} \sim \sqrt{8/\pi n}$, as $n \to \infty$. Therefore we travel from $(2,2)$ to $(0,2)$ much faster than from $(0,2)$ to $(0,1)$ and, given $V_1 = (2,2)$, we again have $\mathbb{P}\{\widehat{T}^{(3)}(1) > n\} \sim c/n^{1/4}$, as $n \to \infty$.

Finally, let $V_1 = (2,0)$ (Figure 1b). From the equation (125) we have $V_2 \stackrel{d}{=} (2 + \xi_2^{(1)}, 1)$ and $V_{\widehat{T}^{(3)}(1)} = (0,1)$, where $\mathbb{P}\{\widehat{T}^{(3)}(1) > n\} \sim \sqrt{8/\pi n}$, as $n \to \infty$.

Thus,

$$\mathbb{P}\{\widehat{T}^{(3)}(1) > n\} \sim \frac{1}{2}\mathbb{P}\{\sum_{k=1}^{\tau} \tau_k > n\} \sim \frac{c}{2n^{1/4}}, \text{ as } n \to \infty. \quad (129)$$

□

Thus, we get the relation between time-instants $T^{(3)}(1)$ and $\{\widehat{T}^{(3)}(k)\}_{k=1}^{\infty}$. Each time we are at the auxiliary state we have a probability $1/4$ to jump into the state

$D(n) = C(n) = M(n)$ independent of anything else. Using the latter lemma and the results of Section 1.5 from Borovkov & Borovkov (2008) we get the following result.

Proposition 4.2. *Let* $\nu = \inf\{k \geq 1 : \widehat{T}^{(3)}(k) - \widehat{T}^{(3)}(k-1) = 1\}$. *Then* ν *has a geometric distribution with parameter* $1/4$ *and*

$$\mathbb{P}\{J_1^{(3)} > n\} = \mathbb{P}\{\widehat{T}^{(3)}(\nu) > n\} \sim 4\mathbb{P}\{\widehat{T}^{(3)}(1) > n\}, \ \textit{as } n \to \infty \tag{130}$$

and therefore there exists a positive random variable $D^{(3)}$ *with a stable distribution such that*

$$\frac{T^{(3)}(n)}{n^4} = \frac{\sum_{k=1}^{n} J_k^{(3)}}{n^4} \Rightarrow D^{(3)}, \ \textit{as } n \to \infty. \tag{131}$$

§ 4.4. Distribution of random variable $Y_1^{(3)}$

In the previous Subsection we analysed the time our process spends between auxiliary states. In this Subsection we analyse the total jumps of the mouse between the states (it can have either zero jumps, one jump, or two jumps).

Let $\{Z_k\}_{k=0}^{\infty}$ be an auxiliary Markov chain which satisfies $Z_k = C(\widehat{T}^{(3)}(k)) - M(\widehat{T}^{(3)}(k)) \in \{-1, 0, 1\}$. Note that $Z_0 = Z_\nu = 0$. Straightforward calculations show that the transition matrix of our auxiliary Markov chain (state "0" is in the beginning, then "1" and then "-1") is

$$P_Z = \begin{pmatrix} \frac{1}{4} & \frac{3}{8} & \frac{3}{8} \\ \frac{1}{4} & \frac{5}{8} & \frac{1}{8} \\ \frac{1}{4} & \frac{1}{8} & \frac{5}{8} \end{pmatrix}. \tag{132}$$

Let

$$\widehat{\xi}_k^{(3)} = M(\widehat{T}^{(3)}(k)) - M(\widehat{T}^{(3)}(k-1)) \text{ for } k \geq 1. \tag{133}$$

Let $\{\psi_k^1\}_{k=1}^{\infty}$ be independent copies of $\tau_1^{(1)}$ and let $\{\psi_k^2\}_{k=1}^{\infty}$ be independent copies of $\sum_{k=1}^{\tau_1^{(1)}} \psi_k^1$. Assume that ζ has binomial distribution with parameter $1/2$ independent of $\{\psi_k^1\}_{k=1}^{\infty}$ and $\{\psi_k^2\}_{k=1}^{\infty}$. Then we have

$$\mathbb{P}\{\widehat{\xi}_2^{(3)} = \pm 1, \ \widehat{T}^{(3)}(2) - \widehat{T}^{(3)}(1) = 1 |\ Z_1 = 0, Z_2 = 0\} = \frac{1}{2}, \tag{134}$$

$$\mathbb{P}\{\widehat{\xi}_2^{(3)} = 0, \ \widehat{T}^{(3)}(2) - \widehat{T}^{(3)}(1) = 1 |\ Z_1 = \pm 1, Z_2 = 0\} = 1, \tag{135}$$

$$\mathbb{P}\{\widehat{\xi}_2^{(3)} = m,\ \widehat{T}^{(3)}(2) - \widehat{T}^{(3)}(1) = k |\ Z_1 = 0, Z_2 = \pm 1\}$$

$$= \begin{cases} \frac{1}{6}\mathbb{P}\{\psi_2^1 = k\}, & \text{if } m = \pm 2 \text{ or } m = 0, \\ \frac{2}{3}\mathbb{P}\{\zeta\psi_2^1 + \psi_2^2 = k\}, & \text{if } m = \pm 1, \end{cases} \tag{136}$$

$$\mathbb{P}\{\widehat{\xi}_2^{(3)} = m,\ \widehat{T}^{(3)}(2) - \widehat{T}^{(3)}(1) = k |\ Z_1 = \pm 1, Z_2 = \pm 1\}$$

$$= \begin{cases} \frac{1}{5}\mathbb{P}\{\psi_1 = k\}, & \text{if } m = \pm 1, \\ \frac{4}{5}\mathbb{P}\{\zeta\psi_1 + \psi_2 = k\}, & \text{if } m = 0, \end{cases} \tag{137}$$

$$\mathbb{P}\{\widehat{\xi}_2^{(3)} = \mp 1, \widehat{T}^{(3)}(2) - \widehat{T}^{(3)}(1) = k |\ Z_1 = \pm 1, Z_2 = \mp 1\} = \mathbb{P}\{\psi_1 = k\}. \tag{138}$$

We have $|\widehat{\xi}_1^{(3)}| \le 2$ and random variable ν has a light-tailed distribution (see Subsection **1.3**). Therefore, random variable $Y_1^{(3)} = M(J_1^{(3)}) = \sum_{k=1}^{\nu} \widehat{\xi}_k^{(3)}$ has a light-tailed distribution. Using that and a symmetry argument we get the following result.

Proposition 4.3. *We have* $\mathbb{E}Y_1^{(3)} = 0$ *and* $\mathbb{E}(Y_1^{(3)})^m < \infty$, *for any* $m \ge 2$.

§ 4.5. Proof of Theorem 1.21

Random vectors $\{Y_n^{(3)}, J_n^{(3)}\}_{n=1}^{\infty}$ are independent and identically distributed, where $Y_1^{(3)} = \sum_{k=1}^{\nu} \widehat{\xi}_k^{(3)}$ and $J_1^{(3)} = T^{(3)}(1) = \widehat{T}^{(3)}(\nu)$. We have

$$\eta(t) = \max\{n > 0 :\ \sum_{k=1}^{n} J_k^{(3)} \le t\} \text{ and } \widetilde{M}(t) = \sum_{k=1}^{\eta(t)} Y_k^{(3)}. \tag{139}$$

From Propositions **4.2** and **4.3** we have

$$\mathbb{E}Y_1^{(3)} = 0,\ \mathbb{E}(Y_1^{(3)})^m < \infty,\ \text{for } m \ge 2, \text{ and } \mathbb{P}\{J_1^{(3)} > n\} \sim \frac{c}{n^{1/4}}, \tag{140}$$

as $n \to \infty$. From Theorem 5.1 from Kasahara (1984) (see also Appendix **A.2**) we have

$$\left\{ \frac{\widetilde{M}(nt)}{\sqrt{\mathbf{Var}Y_1^{(3)}}\, n^{1/8}}, t \ge 0 \right\} \overset{\mathcal{D}}{\Rightarrow} \{B(E^{(3)}(t)), t \ge 0\},\ \text{as } n \to \infty, \tag{141}$$

where $B(t)$ is a standard Brownian motion, independent of $E^{(3)}(t)$.

We show now that (141) holds with $M(nt)$ in the place of $\widetilde{M}(nt)$. It is sufficient to prove that for any fixed $T > 0$

$$\frac{\max_{1\le k\le [nT]}\left\{\widetilde{M}_k - M_k\right\}}{n^{1/8}} \stackrel{a.s.}{\to} 0 \text{ as } n \to \infty. \tag{142}$$

In the time interval $(\eta(nt), nT]$ there are no time-instants n when $D(n) = C(n) = M(n)$, however the mouse may have jumps. Nevertheless, the number of this jumps can be bounded by $2\widehat{\nu}$, where the random variable $\widehat{\nu}$ has a geometric distribution with parameter $1/4$.

Let $U_n = \max_{T^{(3)}(n-1)\le l\le T^{(3)}(n)} |\widetilde{M}_l - M_l|$, $n \ge 1$. We have

$$\max_{1\le l\le J_1^{(3)}} |\widetilde{M}_l - M_l| \stackrel{d}{\le} \sum_{k=1}^{\nu} |\xi_k|. \tag{143}$$

Proposition 4.4. *For any $m \ge 1$ we have $\mathbb{E}U_1^m < \infty$ and $n^{-1/m}\max_{1\le l\le n} U_l$ converges to 0 a.s., as $n \to \infty$.*

We have $\eta(nT) \to \infty$ a.s. and there exists an random variable ζ such that $n^{-1/4}\eta(nT) \Rightarrow \zeta$, as $n \to \infty$ (see, e.g., Section XI.5 in Feller (1971b) or Subsection **1.3.3**). Thus,

$$\begin{aligned}\frac{|\max_{1\le k\le [nT]}\left\{\widetilde{M}_k - M_k\right\}|}{n^{1/8}} &\le \frac{\max_{1\le l\le \eta(nT)}\{U_l\}}{n^{1/8}} + \frac{2\widehat{\nu}}{n^{1/8}} \\ &= \frac{\max_{1\le l\le \eta(nT)}\{U_l\}}{\eta^{1/4}(nT)}\left(\frac{\eta(nT)}{n^{1/2}}\right)^{1/4} + \frac{2\widehat{\nu}}{n^{1/8}} \stackrel{a.s.}{\to} 0.\end{aligned} \tag{144}$$

This completes the proof of Theorem **1.21**.

Part II. Stability of neural networks

SECTION 5
Introduction and main results

§ 5.1. Modelling of neural networks

We analyse the stochastic stability of a model of a neural network. Our model is inspired by the *stochastic integrate-and-fire neuron* model. The original model of membrane potentials was introduced by Lapicque (1907) and has been developed over the years (for a review of the model see, e.g., Burkitt (2006a,b)). In this model, at any time t, the internal state of a neuron i is given by its membrane potential $Z_i(t)$, which evolves according to a stochastic differential equation

$$dZ_i(t) = F(Z_i(t), I(t), t)dt + \sigma(Z_i(t), I(t), t)dW_i(t), \tag{145}$$

where F is a drift function, σ the diffusion coefficient, I is the neuronal input, and W_i is a Brownian motion (see, e.g., Gerstner and Kistler (2002)). The process $W_i(t)$ represents combined internal and external noise. The process I models firings of neurons' potentials (or "spikes"): whenever a potential $Z_i(t)$ reaches certain threshold, it resets to a base-level, and the neuron sends signals to other neurons.

A large number of experiments have given us an understanding of the dynamics of a single neuron. For example, Hodgkin and Huxley (1952) found three different types of ion current flowing through a neuron's membrane, and introduced a detailed model of a membrane potential. To give a basic description, without any input, the neuron is at rest, corresponding to a constant membrane potential. Given a small change, the membrane potential returns to the resting position. If the membrane potential is given a big enough increase, it reaches a certain threshold, and exhibits a pulse-like excursion that will affect connected neurons. After the pulse, the membrane potential does not directly return to the resting potential, but goes below it. This is connected to the fact that a neuron can not have two spikes one right after another.

A neuron network may be thought of as a connected graph of neurons with synapses serving as edges between vertices. When a presynaptic neuron fires a spike it sends a signal through a synapse to a postsynaptic neuron. A neuron is called *inhibitory* if its signals predominantly move the membrane potentials away from a threshold; and *excitatory* if they predominantly move potentials toward a threshold. In this paper, we consider a model containing inhibitory neurons only. It is important to point out that the effect of a signal depends on the potential of a receiving neuron. For example, if the membrane potential of a postsynaptic neuron is lower than that of a corresponding inhibitory synapse, the effect of a signal will be reversed. Therefore, in the models where signals and potentials are assumed to be independent, it is important to assume that the potentials should not decrease too much.

A classical model, introduced by Stein (1965), is the so-called *leaky integrate-and-fire neuron* model, where

$$F(Z_i(t), I(t), t) = -\alpha Z_i(t) + I(t), \quad \sigma(Z_i(t), I(t), t) = \sigma = \text{const.} \tag{146}$$

There are several variations of this model. For instance, nonlinear models were considered, such as the quadratic model (see, e.g., Latham *et al.* (2000)) where $-\alpha Z_i(t)$ is replaced with $a(Z_i(t) - z_{rest})(Z_i(t) - z_c)$, where $z_c > z_{rest}$. Another direction for generalisation of this model is the Spike Response Model (see, e.g., Gerstner and Kistler (2002), Chapter 4.2). In this model, the relation between the dynamics and the potential is determined by the time of the last spike. This allows one to explicitly forbid spikes to occur one right after another and to write the dynamics in integrated form.

5.1.1. Perfect integrate-and-fire neuron model

In the thesis we consider the so-called *perfect integrate-and-fire neuron* model, where $\alpha = 0$ and, therefore, the decay of the membrane potential over time is neglected. This restriction is a stepping stone to achieve more general results and it allows us to write the model in integrated form.

In our model, the spikes and corresponding signals are represented by shifts from a threshold of a random length, independent of everything else. We analyse the system under certain conditions on the distribution of those shifts and prove stability. Instead of considering the recurrence of sets $[-k, H]^N$ (where H is the threshold and N is the number of neurons), we move each coordinate down and reflect the system to work with more convenient sets $[0, k+H]^N$. Thus, in our model, membrane potentials are nonnegative processes that jump to a random positive level after reaching zero. Signals from inhibitory neurons push membrane potentials away from the threshold, i.e. they are positive shifts. It is important to note that we assume that the travel time of signals between neurons is zero, which in general can cause uncertainty in the order of spikes. However, the inhibitory signals do not cause spikes right away, and we assume that the potentials $Z_i(t)$ almost surely do not reach their thresholds at the same time. We refer to Taillefumier *et al.* (2012) for a further discussion.

It is often assumed that the studied system of neurons is itself a part of a much larger system of neurons. The effect of the larger system on the sub-system under consideration is often modelled by a multivariate Brownian motion $W(t)$ with a drift (the drift guaranties the stability of a system of a single neuron). However, we can generalise it to a multivariate spectrally positive (i.e. with positive jumps only) Lévy process $X(t)$ to account for inhibitory signals. It is important for our analysis that the signals do not influence the dynamics of the process $Z(t)$ if it is away from the threshold, i.e. we have $dZ(t) = dX(t)$ if $Z_i(t) > 0$, for $i \in \{1, \ldots, N\}$. Nevertheless, the number of spikes $\eta_i(t)$ before time t is essential to the stability analysis. The fact that $\eta_i(t)$ is not pathwise monotone with respect to signal sizes or the initial state brings certain difficulties in proving stability.

As mentioned above, the system of a single neuron is stable. However, for a general distribution of signals between neurons, 'partial stability' (see, e.g., Cottrell (1992) and Fricker *et al.* (1994)) can occur when only a (possibly random) subset of neurons stabilises, while membrane potentials of other neurons are "pushed" to infinity (which contradicts the physical setup). This situation is of independent mathematical interest and there is no detailed discussion in the Thesis.

It is possible to analyse this model similarly to queueing networks (see, e.g., Asmussen and Turova (1998) and references therein). For example, instead of membrane potentials one may consider residual times $(R_1(t), \dots, R_N(t))$ until a spike. In other words, $R_i(t)$ is the time that the process $(X_i(t+s) - X_i(t),\ s \geq 0)$ needs to reach $-Z_i(t)$. Thus, we have N queues with workloads decreasing at unit speed. When one of the residual times $R_i(t-)$ becomes zero this triggers a spike, and, potentially, every queue (depends on the connectivity of the system) has an arrival of additional workload. More precisely, assume that $R_i(t-) = 0$ and let ξ_{ij} represent the signals, produced by the spike. Then every residual time $R_j(t-)$ is increased by θ_{ij} which is the time that the process

$$(X_j(t + R_j(t-) + s) - X_j(t + R_j(t-)),\ s \geq 0) \tag{147}$$

needs to reach $-\xi_{ij}$. It is very significant that the arrivals of workloads are synchronised, and it brings interesting consequences for the analysis.

§ 5.2. General overview of our results

The ideas and methods for stability analysis of stochastic systems and networks using scaling limits that are linear both in space and in time have become popular and have been developed in 80's–90's of the last century, thanks to works by V.A. Malyshev and his co-authors (see Malyshev (1972), Malyshev and Menshikov (1979), Ignatyuk and Malyshev (1991), and Malyshev (1993)) where the so-called *second vector filed* has been introduced, and later works by A.N. Rybko, A.L. Stolyar and J. Dai (see Rybko and Stolyar (1992), Dai (1995), and Stolyar (1995)) who introduced *fluid limits*. In our analysis of neural networks, we follow the latter approach. Although this method is usually applied to queueing networks, it is quite universal, and turns out to also be applicable to our model. We also refer to Foss and Konstantopoulos (2004) for an overview of some stochastic stability methods.

In particular, we introduce fluid limits and prove their piecewise linearity under specific conditions on average signals and the drift $\mathbb{E}X(1)$. Then we study convergence of the fluid limits to zero and apply (a version of) the stability criterion introduced by Dai (1995) that says that stability of all fluid limits implies positive recurrence of the underlying Markov process. We then prove the existence of a so-called *minorant measure*. Thus, we can use results from Section 7 of Borovkov and Foss (1992) (see, also, Chapter VII of Asmussen (2003)) to prove convergence to the stationary distribution in total variation.

5.2.1. Fluid model and positive recurrence

Let us start with an example of multiclass queueing network. Let $\{1,\ldots,K\}$ be a set of customer classes and $\{1,\ldots,J\}$ a set of stations. Each station j is a single-server service facility that serves customers from the set of classes $c(j)$ according to a non-idling, work-conserving, non-preemptive, but otherwise general, service discipline. It is assumed that $c(j)\cap c(i)=0$ if $i\neq j$. There is a single arrival stream $A(t)$ of intensity λ. The service for customers from class k goes with intensity μ_k. Routing at the arrival point is done according to probabilities p_k, so that an arriving customer becomes of class k with probability p_k. Routing in the network is done so that a customer finishing service from class k joins class l with probability $p_{k,l}$. Let $A_k(t)$ be the cumulative arrival process of class k customers from the outside world. Let $D_k(t)$ be the cumulative departure process from class k. The process $D_k(t)$ counts the total number of departures from class k, both those that are recycled within the network and those who leave it. Of course, it is the specific service policies that will determine $D_k(t)$ for all k. If we introduce independent identically distributed routing variables $\{\alpha_k(n), n\in\mathbb{Z}^+\}$ so that $\mathbb{P}\{\alpha_k(n)=l\}=p_{k,l}$, then we may write the class-k dynamics as:

$$Q_k(t)=Q_k(0)+A_k(t)+\sum_{l=1}^{K}\sum_{n=1}^{D_l(t)}\mathbb{I}(\alpha_l(n)=k)-D_k(t). \tag{148}$$

There are various other equations satisfied by the queueing system, which we omit.

Now, denote the sum of all initial conditions (initial workload, the time until the first arrival and remaining service times) as N. While it is clear that $A_k(Nt)/N$ has a limit as $N\to\infty$, it is not clear at all that so do $D_k(Nt)/N$. The latter depends on the service policies, and, even if a limit exists, it may exist only along a certain subsequence. We say that $X(\cdot)$ is a limit point of $X_N(\cdot)$ if there exists a deterministic subsequence $\{N_l\}$, such that, $X_{N_l}\to X$, as $l\to\infty$, almost surely and uniformly on every compact set. Denote $D(t)=(D_1(t),\ldots,D_K(t))$. A *fluid limit* is any limit point of the sequence of functions $\{D(Nt)/N, t\geq 0\}$ and the *fluid model* is the set of these limit points. If $\overline{D}(t)=(\overline{D}_1(t),\ldots,\overline{D}_K(t))$ is a fluid limit, then we can define

$$\overline{Q}_k(t)=\overline{Q}_k(0)+\overline{A}_k(t)+\sum_{l=1}^{K}\overline{D}_l(t)p_{l,k}-\overline{D}_k(t),\ \text{for } k=1,\ldots,K. \tag{149}$$

It can be interpreted the following way: since $D(Nt)/t \to \overline{D}(t)$, along, possibly, a subsequence, then, along the same subsequence, $Q(Nt)/N \to \overline{Q}(t)$. This follows from the FLLN for the arrival process and for the switching process.

We say that the fluid model is *stable*, if there exists a deterministic $t_0 > 0$, such that, for all fluid limits, $\overline{Q}(t) = 0$ for $t \geq t_0$, almost surely. Equivalently the fluid model is stable if there exist a deterministic time $t_0 > 0$ and a number $\varepsilon \in (0, 1)$ such that, for all fluid limits, $\overline{Q}(t_0) \leq 1 - \varepsilon$, almost surely.

Proposition 5.1. *Denote the state of the system at the arrival epochs as* $\{X_n\}_{n=1}^{\infty}$*. If the fluid model is stable, then there exists a bounded set* B *which is positive recurrent for* $\{X_n\}_{n=1}^{\infty}$*.*

The latter is proven via Lyapunov function criterion (see Foss and Konstantopoulos (2004)).

The definition of stability of a fluid model is quite a strong one. Nevertheless, if it holds - and it does in many important examples - then the original multiclass network is stable. However, in a case of random fluid limits one might consider alternative weaker versions.

In the thesis we take Markov process $Z^{\mathbf{z}}(t)$ with $Z^{\mathbf{z}}(0) = \mathbf{z} \in \mathbb{R}^{+N}$. Let $\|\mathbf{z}\| = \sum_{i=1}^{N} z_i$ for $z \in \mathbb{R}^{+N}$. We introduce a family of scaled processes

$$\widehat{Z}^{\mathbf{z}} = \left\{ \widehat{Z}^{\mathbf{z}}(t) = \frac{Z^{\mathbf{z}}(\|\mathbf{z}\| t)}{\|\mathbf{z}\|},\ t \geq 0 \right\}. \tag{150}$$

We then analyse the convergence of such processes as $\|z\| \to \infty$. We prove that the fluid limits are piecewise deterministic and linear. Next, we prove stability of all fluid limits, which implies existence of a bounded positive recurrent set. Now, in the example we consider a Markov chain constructed on the arrival epochs. Following the lines of Dai (1995) one can get the result for continuous time. Thus, we prove the existence of a bounded set V such that

$$\tau^{\mathbf{z}}(\varepsilon, V) = \inf\{t \geq \varepsilon : Z^{\mathbf{z}}(t) \in V\} < \infty \text{ almost surely and } \sup_{z \in V} \mathbb{E}\tau^{\mathbf{z}}(\varepsilon, V) < \infty. \tag{151}$$

5.2.2. Ergodicity of processes admitting embedded Markov chains

We discussed positive recurrence for Markov processes. Now we move to ergodicity or convergence to stationary distribution. For reference see, e.g., Borovkov and Foss (1992). Let $(\mathbf{X}, \mathbf{B_X})$ be an arbitrary measurable space and let $X = (X(x,n))_{n=0}^{\infty}$ be an $\mathbf{X}$-valued Markov chain with the initial state $X(x,0) = x \in \mathbf{X}$. Assume that there exist a positive recurrent set V, a probability measure φ on $(\mathbf{X}, \mathbf{B_X})$, a number $p \in (0,1)$, and a non-negative integer $m \geq 0$ such that

$$\inf_{x \in V} \mathbb{P}\{X(x, m+1) \in B\} \geq p\varphi(B) \text{ for any } B \in \mathbf{B_X}. \tag{152}$$

Denote $X(\varphi, n)$ as a Markov chain with a random initial value distributed according to φ. Let $\tau_V(\varphi) = \inf\{n \geq 1 : \ X(\varphi, n) \in V\}$. Let $k_1, k_2, \ldots$ be the integer numbers for which $\mathbb{P}\{\tau_V(\varphi) = k_i\} > 0$. Assume that there exists a number $l > 0$ such that the greatest common divisor of the set $(m + k_1 + 1, m + k_2 + 1, \ldots, m + k_l + 1)$ is equal to one (this condition is satisfied if $m = 0$ and $\varphi(V) > 0$). Then the following proposition holds.

Proposition 5.2. *(Theorem 2.1, Borovkov and Foss (1992)) There exists a stationary Markov chain $(X^n)_{n=0}^{\infty}$ with transition probability $\boldsymbol{P}(x, B)$ defined on the same probability space with X, which independent of $X(0)$ and such that for each $x \in \boldsymbol{X}$*

$$\mathbb{P}\{X(x,k) = X^k \textit{ for all } k \geq n\} \to 1, \ \textit{ as } n \to \infty. \tag{153}$$

Relation (153) implies necessarily that the distribution $\pi(B) = \mathbb{P}\{X^0 \in B\}$ of X^0 is an invariant measure:

$$\pi(B) = \int_{\mathbf{X}} \pi(dx)\mathbf{P}(x, B), \ B \in \mathbf{B_X}, \tag{154}$$

and the convergence in total variation occurs

$$\sup_{B \in \mathbf{B_X}} |\mathbb{P}\{X(x,n) \in B\} - \pi(B)| \to 0, \text{ as } n \to \infty. \tag{155}$$

Such results for Markov chains can help in the analysis of special stochastic processes. Take a stochastic process $\{Z(t), \ t \geq 0\}$. Denote a sequence of embedded times $(T_n)_{n=0}^{\infty}$ such that $T_0 \geq 0$ and $(e_n)_{n=0}^{\infty} = (T_{n+1} - T_n)_{n=0}^{\infty}$ are positive random variables. Denote σ-algebra $F_{(t)} = \sigma\{Z(u); \ u \leq t\}$.

Definition 5.3. Process Z *admits an embedded Markov chain* if there exists a sequence of Markov times $(T_n)_{n=0}^{\infty}$ such that

- the sequence $X(n) = Z(T_n)$ constitutes a homogeneous Markov chain;

- for any $n \geq 0$, $t \geq 0$ the joint distribution of $\{Z(T_n + t),\ (e_{n+k})_{k=0}^{\infty}\}$ depends on $Z(T_n) = X(n)$ and t only, i.e.,

$$\begin{aligned} &\mathbb{P}\{Z(T_n + t) \in B,\ (e_{n+k})_{k=0}^{\infty} \in D|\ F_{(T_n)}\} \\ &= \mathbb{P}\{Z(T_n + t) \in B,\ (e_{n+k})_{k=0}^{\infty} \in D|\ \sigma(X(n))\} \end{aligned} \tag{156}$$

almost surely for any $B \in \mathbf{B_X}$ and $D \in \mathbf{B}_{\mathbf{R}_+^{\infty}}$.

Assume that $X(n)$ follows the same conditions as before (positive recurrent set, minorant measure and aperiodicity). For embedded times we assume

$$\sup_{x \in \mathbf{X}} \mathbb{E}\,(e_0|\ X(0) = x) < \infty. \tag{157}$$

Denote by $\tau(y)$ a random variable with the distribution

$$\mathbb{P}\{\tau(y) > t\} = \mathbb{P}\{T_\mu - T_0 > t|\ X(0) \in dy\}, \tag{158}$$

where μ is the first hitting time of the set V by $X(n)$ and φ is the minorant measure.

Proposition 5.4. *(Theorem 7.3, Borovkov and Foss (1992)) Assume that process $\{Z(t),\ t \geq 0\}$ satisfies the following:*

1. *the distribution*

$$\mathbb{P}\{\widehat{\tau} > t\} = \int \varphi(dy)\mathbb{P}\{\tau(y) > t\} \tag{159}$$

of the random variable $\widehat{\tau}$ is a non-lattice one;

2. *the trajectories of the process Z are right (or left) continuous.*

Then the distributions $\boldsymbol{P}_t(\cdot) = \mathbb{P}\{Z(t) \in \cdot\}$ weakly converge to some probability distribution. If instead of the first condition we require the distribution of $\widehat{\tau}$ to possess an absolutely continuous component, the second condition becomes redundant and the convergence is in total variation.

Thus, to prove convergence in total variation we find a proper embedding (we assume that $(e_n)_{n=0}^{\infty}$ are independent uniformly distributed bewteen one and two) with absolutely continuous component. Through fluid approximation model we prove positive recurrence. Finally, we prove that Lebesgue measure on a bounded set is a minorant measure for embedded Markov chain.

5.2.3. Related results

Let us return to the residual-time model $(R_1(t), \ldots, R_N(t))$. There is a series of papers studying ergodic properties of such models. Assume that $\theta_{ij} = \theta^{(i)}$ for $j \neq i$. Cottrell (1992) study conditions for the model to be ergodic when 'residual times' θ_{ii} are exponential and $\theta^{(i)}$ are constants. Fricker *et al.* (1994) maintained the exponential assumption for θ_{ii} but allowed $\theta^{(i)}$ to have a general distribution. The authors give broad stability analysis for several types of networks. For the embedded Markov chain R_n, which takes the values of $R(t)$ right after the spikes, the authors acquire an explicit form of the stationary distribution. The result involves the stationary waiting times of N separate $M/G/1$ queues, with inter-arrival times and service times for the i-th queue having the same distribution as θ_{ii} and $\theta^{(i)}$ respectively, and it is proved via Laplace transforms. Karpelevich *et al.* (1995) consider a more general architecture of the connections in the network and give necessary and sufficient conditions for ergodicity. Asmussen and Turova (1998) get rid of the exponential assumption on θ_{ii} via a sample path approach which highlights the connection to other applied probability areas such as renewal theory, queueing theory and point processes. This approach allows the authors to generalise the result of Fricker *et al.* (1994) for R_n, and it provides the key for studying further aspects of the neural model.

Next, we go back to the membrane potentials. The *stochastic integrate-and-fire neuron* model has received an increasing amount of attention in recent years. There are a number of papers considering mean-field limits of such systems, i.e., letting the number of neurons N becomes large while assuming interactions of order $1/N$. The resulting nonstandard equations of McKean–Vlasov-type is the subject of Cáceres *et al.* (2011) from a PDE perspective and Delarue *et al.* (2015) from a stochastic point of view. The introduction of excitatory neurons may lead to a very interesting behaviour in the limit. Namely, these works show that, even though for a certain choice of parameters the limit equation has a unique global-in-time solution, nevertheless, there is a choice of parameters which leads to a *blow-up phenomenon* in finite time. In other words, the effect of a single neuron spiking may cause an instantaneous cascade during which a significant portion of neurons has a spike. This brings challenges for the approximation of the finite system by the limiting one after a blow-up.

It is important to mention that in the thesis we consider a relatively simple model as a cornerstone for future research (including mean-field limits and introduction of excitatory neurons). Nevertheless, there is a vast variety of advanced models for which significant results are obtained. Next, we introduce several examples of works dealing with such models.

De Masi *et al.* (2015) consider a special case of N identical inhibitory neurons with $N \to \infty$. Firstly, the evolution of membrane potentials is deterministic between spikes. Secondly, each membrane potential has a drift to the average potential. Finally, spikes do not occur at the moments when a potential reaches a certain threshold. Instead, there is a spiking rate of the system which depends on the current membrane potentials. The spiking neuron returns to the resting potential and sends signals of size $1/N$ to every other neuron. The authors regard the state of the neurons $U^N(t) = (U_1^N(t), \ldots, U_N^N(t))$ as a distribution of $1/N$ valued Dirac masses placed at the positions $U_1^N(t), \ldots, U_N^N(t)$. As N goes to infinity, the authors study the limiting fraction of neurons with potentials belonging to a certain interval and introduce non-linear PDE which describes it. The usual approach to prove hydrodynamic limits in mean-field systems is to show that propagation of chaos holds. However, the potentials are correlated due to instantaneous signals through the whole system. To get the result the authors use coupling methods on a discrete-time version of the process.

Robert and Touboul (2016) also consider a model without a fixed spiking threshold. The authors consider a system of N excitatory neurons. Here spikes occur as a inhomogeneous Poisson process and spiking rate is, again, a function of a current membrane potential. The signals between neurons $\xi_{ij}^{(n)}$ are independent identically distributed random variables. The main difference is the leaking of membrane potentials: without interactions between neurons a potential behaves according to $dZ_i(t) = -Z_i(t)dt$. The authors analyse invariant distributions for finite-sized networks and then for the averaged limit. For the finite case the authors prove that, in the presence of an external noise (spiking rate at 0 is positive), there exists a unique non-trivial invariant distribution for the system, and the system "dies out" (no spikes occur after some time) if the spiking rate at 0 is zero. Additionally, the authors prove Harris positive recurrence using a regeneration argument. For the infinite case the authors prove that, under some growth conditions on the spiking rate and other technical conditions, a mean-field result holds. It is important to mention that discontinuities produced by Poisson processes bring a big challenge for the analysis of this model.

Inglis and Talay (2015) also consider a system of excitatory neurons. However, the authors modify the model in order to deal with the aforementioned blow-up phenomenon. More precisely, the signals are constant, however, instead of instantaneous transmission between neurons the influence of a spike is described by a cable equation: if a neuron i spikes at time s then the total impact from that spike on neuron j at time $t > s$ is $J_{ij} \int_s^t G(t-s)ds$. The authors investigate convergence of the averaged system, and they prove the existence and uniqueness of solution for the limiting system.

§ 5.3. Structure of Part II

This part of the Thesis is structured as follows. In Subsection **5.4** we define our model, introduce auxiliary concepts and notations, and formulate our results. In Section **6** we prove Theorem **5.6**. In particular, in Subsection **6.1** we introduce the fluid model and formulate related technical results. In Subsection **6.2** we prove important auxiliary results. In Subsection **6.4** we discuss possible generalisations of our results. In Subsection **6.3** we prove positive recurrence. In Subsection **6.5** we prove that our model satisfies the classical ”minorization” condition. In Section **7** we consider two simple examples of our model, show possible characteristics we want to acquire in general setting, and introduce another direction for future research. Section **8** is an illustration of various directions for the future research. Appendix **B** includes the remaining auxiliary results and comments.

§ 5.4. Model and results

We analyse a network of N stochastic perfect integrate-and-fire inhibitory neurons. At any time t, the internal state of all neurons is given by a multidimensional process $Z(t)$ which represents neurons' membrane potential. Let $X(t)$ be a N-dimensional spectrally positive left-continuous Lévy process with a finite mean, and assume that its distribution has a non-degenerate absolute continuous component. The process $X(t)$ represents combined internal and external noise. Let $\mu_i = -\mathbb{E}X_i(1) > 0$ and $X_i^0(t) = \mu_i t + X_i(t)$. While $Z(t) \in (0, \infty)^N$, membrane potentials evolve as the process $X(t)$, i.e. $dZ(t) = dX(t)$. However, if one of the coordinates becomes non-positive, a shift of independent size occurs for all the coordinates. Let us describe this shift. Let $\{\{\xi_{ij}^{(k)}\}_{i,j=1}^N\}_{k=1}^\infty$ be i.i.d. random matrices, independent of everything else, with a.s. strictly positive elements. Let $b_{ij} = \mathbb{E}\xi_{ij}^{(1)} < \infty$ and $S_{ij}(n) = \sum_{k=1}^n \xi_{ij}^{(k)}$, for $i, j \in \{1, 2, \ldots, N\}$. If potential $Z_i(t)$ hits non-positive values for the k-th time, then instantaneously it increases to $\xi_{ii}^{(k)}$ and other membrane potentials increase by $\xi_{ij}^{(k)}$. We call this event "a spike of neuron i“.

Thus, we can see the process $Z(t)$ as a 'reflected version' of the process $X(t)$: whenever the trajectory hits the border it bounces away. However, when one coordinate "bounces away" all other coordinates are pushed as well. Therefore, we can say that each surface $z_i = 0$ is assigned to a vector $(b_{i1}, b_{i2}, \ldots, b_{iN})$ which describes an average change of trajectory upon hitting the surface. Such analogy works only because we restrict the trajectories of $X(t)$ to not have jumps downward and signals $\xi_{ij}^{(k)}$ are positive. Thus, the trajectory can not jump over the border and the shift given by a spike is independent of the past.

Remark 5.5. One can allow the absolute continuous component of the distribution of the process $X_i(t)$ to be degenerate (for example, take a sum of a Poisson process and a linear function $-at$) and, instead, assume the distribution of the matrix $\{\xi_{ij}^{(1)}\}_{i,j=1}^N$ to have an absolute continuous component. The main result of this part would still hold and the proof would only need few minor changes.

Let $Z^{\mathbf{z}}(t) = (Z_1^{\mathbf{z}}(t), \ldots, Z_N^{\mathbf{z}}(t)) \in \mathcal{Z} = [0, \infty)^N$ be the membrane potentials at time t with an initial value $\mathbf{z} = (z_1, \ldots, z_N)$. Let $T_{i0}^{\mathbf{z}} = 0$ and let

$$T_{ik}^{\mathbf{z}} = \inf\{t > T_{i(k-1)}^{\mathbf{z}} : \ Z_i^{\mathbf{z}}(t) \leq 0\}, \quad \text{for } k \geq 1, \tag{160}$$

be the times when neuron i reaches its threshold. Let $\eta_i^{\mathbf{z}}(0) = 0$ and let $\eta_i^{\mathbf{z}}(t) = \max\{k : \ T_{ik}^{\mathbf{z}} < t\}$ be the number of spikes of $Z_i^{\mathbf{z}}(t)$ before time t. Then the dynamics of the system is given by

$$Z_i^{\mathbf{z}}(t) = z_i + X_i(t) + \sum_{j=1}^N \sum_{k=1}^{\eta_j^{\mathbf{z}}(t)} \xi_{ji}^{(k)} = z_i + X_i(t) + \sum_{j=1}^N S_{ji}(\eta_j^{\mathbf{z}}(t)), \quad i = 1, \ldots, N. \tag{161}$$

Before talking about stability of the system it is important to point out that, due to the negative drift, one can easily show that the potential of an isolated neuron is stable (this is also a subcase of our main result). Nevertheless, there are examples of parameters μ_i and b_{ij} such that there exists a subset of neurons which, after reaching stability, can "push other neurons to infinity". We do not give any details on such cases of partial stability in the thesis. Next, we introduce some simple assumptions which allow to avoid partial stability and prove an ergodic result. However, these conditions are far from necessary for stability and they serve as a nice example where resulting characteristics have a clear and explicit form. We provide a discussion of possible necessary and sufficient conditions in Subsection **6.4**.

We assume that all potentials have the same drift μ and that signals from neuron $i = 1, \ldots, N$ to all other neurons have the same mean w_i. More precisely, we assume that

$$\mu_i = \mu > 0, \quad b_{ij} = \mathbb{E}\xi_{ij}^{(1)} = w_i > 0 \quad \text{and} \quad b_{ii} = H_i > w_i, \quad \text{for } i = 1, \ldots, N \text{ and } j \neq i. \tag{162}$$

However, we allow the distributions of potentials and of signals to differ.

Theorem 5.6. *Assume condition* (162) *to hold. Then the process* $(Z_1^{\mathbf{z}}(t), \ldots, Z_N^{\mathbf{z}}(t))$ *is Harris positive recurrent: there is a distribution* π *such that*

$$\sup_A |\mathbb{P}\{Z^{\mathbf{z}}(t) \in A\} - \pi(A)| \to 0, \ \textit{as } t \to \infty. \tag{163}$$

Remark 5.7. Given (162), matrix $B = (b_{ij})_{i,j=1}^N$ is invertible and

$$(\mathbf{1}B^{-1})_i = \frac{1}{(H_i - w_i)\left(1 + \sum_{k=1}^N \frac{w_k}{H_k - w_k}\right)}, \quad \text{for } i = 1, \ldots, N, \tag{164}$$

where $\mathbf{1} = (1, \ldots, 1)$. Vector $\mu\mathbf{1}B^{-1}$ represents rates of spikes when stability is achieved. In particular, for large $t > \mu^{-1}(1 + \sum_{k=1}^N w_k/(H_k - w_k))$ and for each sequence $\mathbf{z}_n$, $\|\mathbf{z}_n\| \to \infty$, there exists a subsequence $\mathbf{z}_{n_k}$ such that

$$\frac{\eta^{\mathbf{z}_{n_k}}(\|\mathbf{z}_{n_k}\|(t+\Delta)) - \eta^{\mathbf{z}_{n_k}}(\|\mathbf{z}_{n_k}\|t)}{\|\mathbf{z}_{n_k}\|\Delta} \Rightarrow \mu\mathbf{1}B^{-1}, \ \text{for } \Delta > 0. \tag{165}$$

Here we use vector-norm $\|\mathbf{x}\| = \sum_{i=1}^N |x_i|$.

We prove Theorem **5.6** following two standard steps. For the reader's convenience, we formulate those steps as lemmas. Let $\tau^{\mathbf{z}}(\varepsilon, B) = \inf\{t > \varepsilon : \ Z^{\mathbf{z}}(t) \in B\}$

be the first hitting time of a set B after time ε. The first step is the proof of positive recurrence which we achieve via the fluid approximation method.

Lemma 5.8. *There exists $k_0 > 0$ such that for $V = \{\boldsymbol{z} \in \mathcal{Z} : \|\boldsymbol{z}\| < k_0\}$ we have*

$$\sup_{\boldsymbol{z} \in V} \mathbb{E}\tau^{\boldsymbol{z}}(\varepsilon, V) < \infty. \tag{166}$$

In the second step, we show that our model satisfies the classical 'minorization' condition.

Lemma 5.9. *There exist a number $p > 0$ and a probability measure ψ, such that for a uniformly distributed random variable $U \in [1, 2]$, independent of everything else, we have*

$$\inf_{\boldsymbol{z} \in V} \mathbb{P}\{Z^{\mathbf{z}}(U) \in B\} \geq p\psi(B). \tag{167}$$

Using Lemmas **5.8** and **5.9** we can prove that conditions of Theorem 7.3 from Borovkov and Foss (1992) (see also Subsection **5.2.2**) are satisfied, which implies Harris positive recurrence. More precisely, let $(T_n)_{n=0}^{\infty}$ be a sequence of embedded times. Let $(e_n)_{n=0}^{\infty} = (T_{n+1} - T_n)_{n=0}^{\infty}$ are independent random variables uniformly distributed between one and two. Then the Markov chain $(Z(T_n))_{n=0}^{\infty}$ is Harris positive recurrent. Additionally, measure ψ and the distribution of e_0 follow certain 'spread out property', meaning, they have an absolute continuous component. Thus, we have "enough mixing" in continuous time and process $(Z(t),\ t \geq 0)$ is Harris positive recurrent as well.

The proof of Lemma **5.8** is based on the fluid approximation. We dedicate Subsection **6.1** formulating corresponding definitions and auxiliary results. We point out that we need to assume condition (162) only in the proof of Lemma **5.8** and in Remark **5.7**.

One of the difficulties of our model is the lack of path-wise monotonicity for the number of spikes $\eta^{\mathbf{z}}(t)$ with respect to signals $\xi_{ij}^{(k)}$ or initial state $\mathbf{z}$. In general, making one neuron firing a spike earlier may lead to other spikes occurring later. However, we prove that there is a "partial monotonicity" which allows us to get an upper bound for process $\eta^{\mathbf{z}}(t)$ with useful properties.

Since all neurons are inhibitory, one way to increase the number of spikes is to remove all interactions between neurons. Let the process $\widetilde{Z}^{\mathbf{z}}$ be the transformation of the process $Z^{\mathbf{z}}$ by replacing signals $\xi_{ji}^{(k)}$, $j \neq i$, by 0 for $k \geq 1$ (trajectories of $X(t)$ remain the same). The resulting process has a simpler dependence between coordinates and it has a greater number of spikes before any time $t > 0$ than that of $Z^{\mathbf{z}}$. For our convenience, we want to remove the dependence of the upper bound on $\mathbf{z}$ (which is significant because we take $\mathbf{z}$ large in the following lemmas) and make the time until the first spike to have the same distribution with the rest of waiting times. Let the process $\bar{Z}$ be the transformation of the process $\widetilde{Z}^{\mathbf{z}}$, so that $\bar{Z}_i(0) \stackrel{d}{=} \xi_{ii}^{(1)}$, $1 \leq i \leq N$. Let $\widetilde{\eta}^{\mathbf{z}}$ and $\bar{\eta}$ be the number of spikes in processes $\widetilde{Z}^{\mathbf{z}}$ and $\bar{Z}$, respectively.

Lemma 5.10. *We have*

$$\eta_i^{\mathbf{z}}(t) \leq \widetilde{\eta}_i^{\mathbf{z}}(t), \quad a.s., \tag{168}$$

$$\widetilde{\eta}_i^{\mathbf{z}}(t) \stackrel{st.}{\leq} 1 + \bar{\eta}_i(t), \tag{169}$$

and $\bar{\eta}_i(t)$ is an undelayed renewal process, which satisfies the integral renewal theorem and the SLLN

$$\frac{\mathbb{E}\bar{\eta}_i(t)}{t} \to \frac{\mu_i}{b_{ii}} \stackrel{a.s.}{\leftarrow} \frac{\bar{\eta}_i(t)}{t} \tag{170}$$

In the next Section, we present a stability analysis of a class of neural networks with a general Lévy input. Then, in Section **7**, we present two particular examples of the model, with constant input and with Poisson input, where we can find exact expressions for a number of characteristics.

SECTION 6

Stability of the system with Lévy input

In this Section we introduce the fluid model, prove necessary auxiliary results and, finally, give a proof of Theorem **5.6** divided into two parts. We also make some brief comments on more general conditions for stability of the system in Subsection **6.4**.

§ 6.1. Fluid model and corresponding auxiliary results

Let us define the fluid approximation model. Let $\rho(\mathbf{x},\mathbf{y}) = \sum_{i=1}^{N} |x_i - y_i|$ be the metric on $\mathcal{Z}$ and let $\|\mathbf{x}\| = \rho(\mathbf{x}, 0)$, for $\mathbf{x},\mathbf{y} \in \mathcal{Z}$. For each $\mathbf{z} \in \mathcal{Z}$, introduce a family of scaled processes

$$\widehat{Z}^{\mathbf{z}} = \left\{ \widehat{Z}^{\mathbf{z}}(t) = \frac{Z^{\mathbf{z}}(\|\mathbf{z}\|t)}{\|\mathbf{z}\|},\ t \geq 0 \right\}. \tag{171}$$

We call the family

$$\widehat{Z} = \{\widehat{Z}^{\mathbf{z}},\ \|\mathbf{z}\| \geq 1\} \tag{172}$$

relatively compact (at infinity) if, for each sequence $\widehat{Z}^{\mathbf{z}_n}$, $\|\mathbf{z}_n\| \to \infty$, there exists a subsequence $\widehat{Z}^{\mathbf{z}_{n_k}}$ that converges weakly (in the Skorokhod topology) to some limit process $\varphi^Z = \{\varphi^Z(t),\ t \geq 0\}$, which is called a *fluid limit*. A family of such limits is called a *fluid model*. The fluid model is *stable* if there exists a finite constant T such that $\|\varphi^Z(T)\| = 0$ a.s. for any fluid limit φ^Z (there are several equivalent definitions of stability of a fluid model, see e.g. Stolyar (1995)). Based on stability of a fluid model, one can prove positive recurrence of the original Markov process following the lines of Dai (1995).

Using Lemma **5.10** we prove the next result.

Lemma 6.1. *The family of processes $\{Z^{\mathbf{z}},\ \mathbf{z} \in \mathcal{Z}\}$ is such that*

- *for all $t > 0$ and $\mathbf{z} \in \mathcal{Z}$,*

$$\mathbb{E}\|Z^{\mathbf{z}}(t)\| < \infty \tag{173}$$

and moreover, for any K,

$$\sup_{\|\boldsymbol{z}\|\leq K} \mathbb{E}\|Z^{\mathbf{z}}(t)\| < \infty; \tag{174}$$

- *for all $0 \leq u < t$, the family of random variables*

$$\{\rho(\widehat{Z}^{\mathbf{z}}(u), \widehat{Z}^{\mathbf{z}}(t)),\ \|\boldsymbol{z}\| \geq 1\} \tag{175}$$

is uniformly integrable and there exists a constant C such that

$$\limsup_{\|\boldsymbol{z}\|\to\infty} \mathbb{P}\{\sup_{u',t'\in[u,t]} \rho(\widehat{Z}^{\mathbf{z}}(u'), \widehat{Z}^{\mathbf{z}}(t')) > C(t-u)\} = 0. \tag{176}$$

With this result, one can follow the lines of the proof of Theorem 7.1 from Stolyar (1995) to obtain the following.

Corollary 6.2. *The family of processes $\widehat{Z}$ is relatively compact and every fluid limit φ^Z is an a.s. Lipschitz continuous function with Lipschitz constant $C+1$.*

We know that a Lipschitz continuous function is differentiable. We call time t_0 a *regular point* if $\varphi^Z(t)$ is differentiable at t_0. Furthermore, we have

$$\varphi^Z(t) - \varphi^Z(s) = \int_s^t \frac{d\varphi^Z}{du}(u)du, \quad t > s > 0, \tag{177}$$

where the derivative is arbitrarily defined (for example, it equals zero) outside regular points.

Let $\widehat{\eta}^{\mathbf{z}}(t) = \eta^{\mathbf{z}}(\|\mathbf{z}\|t)/\|\mathbf{z}\|$. Following the lines of the proof of Lemma **6.1**, one can prove similar results for the family $\widehat{\eta} = \{\widehat{\eta}^{\mathbf{z}},\ \|\mathbf{z}\| \geq 1\}$. Denote a fluid limit of $\widehat{\eta}$ as φ^η. If at time t we have $\varphi_i^\eta(t) > 0$, then for a certain sequence $\mathbf{z}_n$ the number of spikes $\eta_i^{\mathbf{z}_n}(\|\mathbf{z}_n\|t)$ becomes large. If additionally, $\eta_i^{\mathbf{z}_n}(\|\mathbf{z}_n\|t)$ converges to infinity a.s., then by the law of large numbers

$$\frac{S_{ij}(\eta_i^{\mathbf{z}_n}(\|\mathbf{z}_n\|t))}{\|\mathbf{z}_n\|} = \frac{\eta_i^{\mathbf{z}_n}(\|\mathbf{z}_n\|t)}{\|\mathbf{z}_n\|}\frac{S_{ij}(\eta_i^{\mathbf{z}_n}(\|\mathbf{z}_n\|t))}{\eta_i^{\mathbf{z}_n}(\|\mathbf{z}_n\|t)} \Rightarrow \varphi_i^\eta(t)b_{ij}, \text{ as } n\to\infty. \tag{178}$$

If $\varphi_i^\eta(t) = 0$, then the number of spikes is not as large and, if we prove that the left-hand side of the last equation converges to zero, the resulting convergence will be of the same form.

Using this idea we get the following result.

L e m m a 6.3. *Let $\widehat{\eta}^{\boldsymbol{z}_n}$ converge weakly to a fluid limit φ^η for a sequence $\boldsymbol{z}_n$, $\|\boldsymbol{z}_n\| \to \infty$ as $n \to \infty$. Then we have weak convergence of processes*

$$\left(\left(\frac{1}{\|\boldsymbol{z}_n\|} \sum_{i=1}^N S_{ij}(\eta_i^{\boldsymbol{z}_n}(\|\boldsymbol{z}_n\| t)) \right)_{j=1}^N, \ t \geq 0 \right) \overset{D}{\Rightarrow} (\varphi^\eta(t) B, \ t \geq 0). \tag{179}$$

§ 6.2. Proofs of auxiliary results

In this Subsection we prove our auxiliary results for a general matrix B and parameters μ_i.

6.2.1 *Proof of Lemma* ***5.10***

We prove that $T^{\mathbf{z}}_{ik} \geq \widetilde{T}^{\mathbf{z}}_{ik}$:

$$\begin{aligned} T^{\mathbf{z}}_{ik} &= \inf\{t > T^{\mathbf{z}}_{i(k-1)} : \ Z^{\mathbf{z}}_i(t) \leq 0\} = \inf\{t > T^{\mathbf{z}}_{i(k-1)} : \ Z^{\mathbf{z}}_i(t) = 0\} \\ &= \inf\{t > T^{\mathbf{z}}_{i(k-1)} : \ z_i + X_i(t) + \sum_{j=1}^N S_{ji}(\eta^{\mathbf{z}}_j(t)) = 0\} \\ &= \inf\{t > T^{\mathbf{z}}_{i(k-1)} : \ z_i + X_i(t) + S_{ii}(k-1) + \sum_{j \neq i} S_{ji}(\eta^{\mathbf{z}}_j(t)) = 0\} \\ &\geq \inf\{t > T^{\mathbf{z}}_{i(k-1)} : \ z_i + X_i(t) + S_{ii}(k-1) = 0\}. \end{aligned} \tag{180}$$

Since $T^{\mathbf{z}}_{i0} = \widetilde{T}^{\mathbf{z}}_{i0} = 0$, by induction we have

$$T^{\mathbf{z}}_{ik} \geq \inf\{t > \widetilde{T}^{\mathbf{z}}_{i(k-1)} : \ z_i + X_i(t) + S_{ii}(k-1) = 0\} = \widetilde{T}^{\mathbf{z}}_{ik}. \tag{181}$$

Thus, we get $\eta^{\mathbf{z}}_i(t) \leq \widetilde{\eta}^{\mathbf{z}}_i(t)$. Since $\widetilde{\eta}^{\mathbf{z}}_i(t) - 1$ has the same distribution as $\bar{\eta}_i(t - \widetilde{T}^{\mathbf{z}}_{i1}) \leq \bar{\eta}_i(t)$, we have the second inequality.

The process $\bar{\eta}_i(t)$ is an undelayed renewal process with waiting times having the same distribution as $\tau_i = \inf\{t > 0 : X_i(t) = -\xi^{(1)}_{ii}\}$. Using the strong law of large numbers, one can prove that $\mathbb{E}\tau_i = b_{ii}/\mu_i$ (see also Borovkov (1965) for a detailed proof). Therefore, via the standard argument of renewal theory the rest of the proof follows (see e.g. Feller (1971)).

6.2.2 *Proof of Lemma* ***6.1***

Part 1. Using Lemma **5.10** and positivity of $\xi^{(k)}_{ij}$, we get

$$\|Z^{\mathbf{z}}(t)\| = \sum_{i=1}^N |z_i + X_i(t) + \sum_{j=1}^N S_{ji}(\eta^{\mathbf{z}}_j(t))| \leq \|\mathbf{z}\| + \|X(t)\| + \sum_{i=1}^N \sum_{j=1}^N S_{ji}(\widetilde{\eta}^{\mathbf{z}}_j(t)). \tag{182}$$

We have

$$\{\widetilde{\eta}_i^{\mathbf{z}}(t) = m\} = \{-\sum_{k=1}^{m} \xi_{ii}^{(k)} < z_i + \inf_{0\le s\le t} X(s) \le -\sum_{k=1}^{m-1} \xi_{ii}^{(k)}\} \tag{183}$$

and, therefore,

$$\widetilde{\eta}_i^{\mathbf{z}}(t) = \inf\{m \in \mathbb{Z}^+ : \sum_{k=1}^{m} \xi_{ii}^{(k)} > -z_i - \inf_{0\le s\le t} X(s)\}. \tag{184}$$

Since $\{\{\xi_{ij}^{(k)}\}_{i,j=1}^N\}_{k=1}^\infty$ and $(X(t),\ t\ge 0)$ are independent, the random variable $\widetilde{\eta}_i^{\mathbf{z}}(t)$ is a stopping time for the sequence $\{\{\xi_{ij}^{(k)}\}_{i,j=1}^N\}_{k=1}^\infty$. By Wald's identity,

$$\mathbb{E}S_{ji}(\widetilde{\eta}_j^{\mathbf{z}}(t)) = \mathbb{E}\sum_{k=1}^{\widetilde{\eta}_j^{\mathbf{z}}(t)} \xi_{ji}^{(k)} = \mathbb{E}\widetilde{\eta}_j^{\mathbf{z}}(t) b_{ji} < \infty. \tag{185}$$

Part 2. We have

$$\begin{aligned}\rho(\widehat{Z}^{\mathbf{z}}(u), \widehat{Z}^{\mathbf{z}}(t)) &= \sum_{i=1}^{N} \frac{|Z_i^{\mathbf{z}}(\|\mathbf{z}\|t) - Z_i^{\mathbf{z}}(\|\mathbf{z}\|u)|}{\|\mathbf{z}\|} \le (t-u)\sum_{i=1}^{N}\mu_i \\ &+ \sum_{i=1}^{N} \frac{|X_i^0(\|\mathbf{z}\|t) - X_i^0(\|\mathbf{z}\|u)|}{\|\mathbf{z}\|} + \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{S_{ji}(\eta_j^{\mathbf{z}}(\|\mathbf{z}\|t)) - S_{ji}(\eta_j^{\mathbf{z}}(\|\mathbf{z}\|u))}{\|\mathbf{z}\|}.\end{aligned} \tag{186}$$

Process X_i is a Lévy process, from which we have

$$\mathbb{E}\frac{|X_i^0(\|\mathbf{z}\|t)|}{\|\mathbf{z}\|} \le 2\sup_{0\le s\le t} \mathbb{E}|X_i^0(s)|, \text{ for } \|\mathbf{z}\| \ge 1, \tag{187}$$

and, therefore, the second summand on the right-hand side of (186) is uniformly integrable. By Lemma **5.10**, we have

$$S_{ij}(\eta_j^{\mathbf{z}}(\|\mathbf{z}\|t)) - S_{ij}(\eta_j^{\mathbf{z}}(\|\mathbf{z}\|u)) \overset{st.}{\le} S_{ij}(1 + \bar{\eta}_i(\|\mathbf{z}\|(t-u))). \tag{188}$$

Since $S_{ij}(n)/n \to b_{ij}$ and $\bar{\eta}_i(\|\mathbf{z}\|(t-u)) \to \infty$ a.s., we have

$$\frac{S_{ij}(1 + \bar{\eta}_i(\|\mathbf{z}\|(t-u)))}{1 + \bar{\eta}_i(\|\mathbf{z}\|(t-u))} \overset{a.s.}{\to} b_{ij}, \tag{189}$$

and therefore

$$\begin{aligned}0 &\le \frac{S_{ji}(\eta_j^{\mathbf{z}}(\|\mathbf{z}\|t)) - S_{ji}(\eta_j^{\mathbf{z}}(\|\mathbf{z}\|u))}{\|\mathbf{z}\|} \overset{st.}{\le} \frac{S_{ij}(1 + \bar{\eta}_i(\|\mathbf{z}\|(t-u)))}{\|\mathbf{z}\|} \\ &= \frac{1 + \bar{\eta}_i(\|\mathbf{z}\|(t-u))}{\|\mathbf{z}\|} \frac{S_{ij}(1 + \bar{\eta}_i(\|\mathbf{z}\|(t-u)))}{1 + \bar{\eta}_i(\|\mathbf{z}\|(t-u))} \to (t-u)\frac{\mu_i}{b_{ii}} b_{ij}\end{aligned} \tag{190}$$

a.s. and in L_1, as $\|\mathbf{z}\| \to \infty$. Then the distance on the left-hand side of (186) is bounded above by the sum of uniformly integrable random variables and therefore is also uniformly integrable.

Given

$$C > \sum_{i=1}^{N} \mu_i \left(1 + \sum_{j=1}^{N} \frac{b_{ij}}{b_{ii}} \right), \tag{191}$$

there exists $\varepsilon > 0$ such that for $\|\mathbf{z}\|$ large

$$\begin{aligned}
&\mathbb{P}\{ \sup_{u',t' \in [u,t]} \rho(\widehat{Z}^{\mathbf{z}}(u'), \widehat{Z}^{\mathbf{z}}(t')) > C(t-u)\} \\
&\leq \mathbb{P}\left\{ \sup_{u',t' \in [u,t]} \sum_{i=1}^{N} \frac{|X_i^0(\|\mathbf{z}\|t') - X_i^0(\|\mathbf{z}\|u')|}{\|\mathbf{z}\|} > \varepsilon(t-u) \right\} \\
&\leq 2N\mathbb{P}\left\{ \sup_{s \in [0,t-u]} \frac{|X_1^0(\|\mathbf{z}\|s)|}{\|\mathbf{z}\|} > \frac{\varepsilon}{2N}(t-u) \right\} \to 0,
\end{aligned} \tag{192}$$

by Theorem 36.8 from Sato (1999).

6.2.3 *Proof of Lemma **6.3***

By Skorokhod (1956) (see also Subsection **1.5**), it is sufficient to prove that there is a convergence of finite-dimensional distributions on everywhere dense set of times t and that a tightness condition holds. Tightness can be deduced from the second statement of Lemma **6.1**. We prove that

$$\mathbb{P}\left\{ \bigcap_{k=1}^{K} \bigcap_{i,j=1}^{N} \left\{ \frac{S_{ij}(\eta_i^{\mathbf{z}_n}(\|\mathbf{z}_n\|t_k))}{\|\mathbf{z}_n\|} < y_{ij}^k \right\} \right\} \to \mathbb{P}\left\{ \bigcap_{k=1}^{K} \bigcap_{i=1}^{N} \left\{ \varphi_i^{\eta}(t_k) < \min_{1 \leq j \leq N} \frac{y_{ij}^k}{b_{ij}} \right\} \right\} \tag{193}$$

as $n \to \infty$, for appropriate $t \geq 0$ and $\mathbf{y} \in (0,\infty)^{KN^2}$.

Define sets

$$C_{ij}^k(n) = \left\{ \frac{S_{ij}(\eta_i^{\mathbf{z}_n}(\|\mathbf{z}_n\|t_k))}{\|\mathbf{z}_n\|} < y_{ij}^k \right\}, \tag{194}$$

$$D_i^k(n,m) = \{\eta_i^{\mathbf{z}_n}(\|\mathbf{z}_n\|t_k) > m\}, \tag{195}$$

$$E_{ij}^k(n,\delta) = \left\{ \left| \frac{S_{ij}(\eta_i^{\mathbf{z}_n}(\|\mathbf{z}_n\|t_k))}{\eta_i^{\mathbf{z}_n}(\|\mathbf{z}_n\|t_k)} - b_{ij} \right| \leq \delta \right\}, \tag{196}$$

$$F_i^{k\pm}(n,\delta) = \left\{ \frac{\eta_i^{\mathbf{z}_n}(\|\mathbf{z}_n\|t_k)}{\|\mathbf{z}_n\|} < \min_{1 \leq j \leq N} \frac{y_{ij}^k}{b_{ij} \mp \delta} \right\}, \tag{197}$$

where $\delta \in (0, \min_{i,j} b_{ij})$. We prove that

$$\mathbb{P}\left\{F_i^{k-}(n,\delta)\right\} + o(1) \leq \mathbb{P}\left\{ \bigcap_{j=1}^{N} C_{ij}^k(n) \right\} \leq \mathbb{P}\left\{F_i^{k+}(n,\delta)\right\} + o(1), \text{ as } n \to \infty. \tag{198}$$

For any $\mathbf{y} \in (\mathbb{R}^+)^{KN^2}$ such that $(\min_j(y^k_{ij}/b_{ij}))^N_{i=1}$ is a continuity point of the cdf of $(\varphi^\eta(t_k))^k_{k=1}$, there is a neighbourhood Δ of $\mathbf{y}$ such that every point $\mathbf{x} \in \Delta$ is also a continuity point. Thus, for δ small we have

$$\begin{aligned}\mathbb{P}\left\{\bigcap_{k=1}^{K}\bigcap_{i=1}^{N} F_i^{k\pm}(n,\delta)\right\} &= \mathbb{P}\left\{\bigcap_{k=1}^{K}\bigcap_{i=1}^{N}\left\{\frac{\eta_i^{\mathbf{z}_n}(\|\mathbf{z}_n\|t_k)}{\|\mathbf{z}_n\|} < \min_{1\le j\le N}\frac{y^k_{ij}}{b_{ij}\mp\delta}\right\}\right\} \\ &\to \mathbb{P}\left\{\bigcap_{k=1}^{K}\bigcap_{i=1}^{N}\left\{\varphi_i^\eta(t_k) < \min_{1\le j\le N}\frac{y^k_{ij}}{b_{ij}\mp\delta}\right\}\right\}, \text{ as } n\to\infty,\end{aligned} \tag{199}$$

and, therefore, by letting δ converge to 0, we get (193).

By the law of large numbers, we have

$$\mathbb{P}\{D_i^k(n,m)\cap\overline{E_{ij}^k(n,\delta)}\}\to 0, \quad \text{as } m\to\infty, \tag{200}$$

and

$$\mathbb{P}\{\overline{C_{ij}^k(n)}\cap\overline{D_i^k(n,m)}\}\to 0, \quad \text{as } n\to\infty, \tag{201}$$

if $m = o(\|\mathbf{z}_n\|)$. Take $m = \sqrt{\|\mathbf{z}_n\|}$.

From the definitions we have

$$\begin{aligned}\left(\bigcap_{j=1}^{N} C_{ij}^k(n)\cap D_i^k(n,m)\cap E_{ij}^k(n,\delta)\right) &\subseteq \left(F_i^{k+}(n,\delta)\cap D_i^k(n,m)\cap E_{ij}^k(n,\delta)\right) \\ &= \left(F_i^{k+}(n,\delta)\cap D_i^k(n,m)\right)\setminus\left(F_i^{k+}(n,\delta)\cap D_i^k(n,m)\cap\overline{E_{ij}^k(n,\delta)}\right)\end{aligned} \tag{202}$$

and

$$\left(F_i^{k+}(n,\delta)\cap D_i^k(n,m)\right) = F_i^{k+}(n,\delta)\setminus\left(F_i^{k+}(n,\delta)\cap\overline{D_i^k(n,m)}\right). \tag{203}$$

Since $m = o(\|\mathbf{z}_n\|)$, we have $F_i^{k+}(n,\delta)\cap\overline{D_i^k(n,m)} = \overline{D_i^k(n,m)}$ for n large. Combining everything, we get

$$\begin{aligned}\mathbb{P}\left\{\bigcap_{j=1}^{N} C_{ij}^k(n)\right\} &= \mathbb{P}\left\{\bigcap_{j=1}^{N} C_{ij}^k(n)\cap D_i^k(n,m)\right\} + \mathbb{P}\left\{\bigcap_{j=1}^{N} C_{ij}^k(n)\cap\overline{D_i^k(n,m)}\right\} \\ &\le \mathbb{P}\left\{F_i^{k+}(n,\delta)\right\} - \mathbb{P}\left\{\overline{D_i^k(n,m)}\right\} + \mathbb{P}\left\{\bigcap_{j=1}^{N} C_{ij}^k(n)\cap\overline{D_i^k(n,m)}\right\} + o(1),\end{aligned} \tag{204}$$

as $n \to \infty$, and

$$\mathbb{P}\left\{\bigcap_{j=1}^{N} C_{ij}^k(n) \cap \overline{D_i^k(n,m)}\right\} - \mathbb{P}\left\{\overline{D_i^k(n,m)}\right\} = \mathbb{P}\left\{\bigcup_{j=1}^{N} \overline{C_{ij}^k(n)} \cap \overline{D_i^k(n,m)}\right\} \to 0, \tag{205}$$

as $n \to \infty$. Following the same lines with replacing a set $F_i^{k+}(n,\delta)$ with a set $F_i^{k-}(n,\delta)$ and relations $\subseteq$ and $\leq$ with relations $\supseteq$ and $\geq$, we get the lower bound.

§ 6.3. Proof of positive recurrence

In this Subsection we prove Lemma **5.8**. For that purpose we show that under condition (162) fluid limits $\varphi^Z(t)$ are deterministic and uniquely defined by initial value $\varphi^Z(0)$. Further, each coordinate of a fluid limit is a continuous piecewise linear function which reaches zero and then remains there.

Let sequence $\mathbf{z}_n$, $\|\mathbf{z}_n\| \to \infty$, be such that

$$\widehat{Z}^{\mathbf{z}_n} \overset{\mathcal{D}}{\Rightarrow} \varphi^Z \text{ and } \widehat{\eta}^{\mathbf{z}_n} \overset{\mathcal{D}}{\Rightarrow} \varphi^\eta. \tag{206}$$

By Corollary 1, function φ^Z is a.s. Lipschitz continuous.

Following the lines of the proof of Lemma **6.3**, one can easily show that

$$\left(\widehat{Z}^{\mathbf{z}_n}(t) - \frac{\mathbf{z}_n}{\|\mathbf{z}_n\|} - \frac{X(\|\mathbf{z}_n\|t)}{\|\mathbf{z}_n\|},\ t \geq 0\right) \overset{\mathcal{D}}{\Rightarrow} (\varphi^Z(t) - \varphi^Z(0) + \mu t\mathbf{1},\ t \geq 0). \tag{207}$$

Now, given (161) and Lemma **6.3**, we have

$$(\varphi^\eta(t)B,\ t \geq 0) \overset{d}{=} (\varphi^Z(t) - \varphi^Z(0) + \mu t\mathbf{1},\ t \geq 0). \tag{208}$$

Note that up to this point we did not need any conditions on matrix B and vector $(\mu_i)_{i=1}^N$ other than $b_{ij} > 0$ and $\mu_i > 0$. Next, we start to use conditions (162).

By Remark **5.7**, the matrix B is invertible and we have

$$(\varphi^\eta(t),\ t \geq 0) \overset{d}{=} \left(\left(\varphi^Z(t) - \varphi^Z(0) + \mu t\mathbf{1}\right) B^{-1},\ t \geq 0\right). \tag{209}$$

Since φ^η is a weak limit, we can assume without loss of generality $\varphi^\eta(t) = \left(\varphi^Z(t) - \varphi^Z(0) + \mu t\mathbf{1}\right) B^{-1}$. Thus, φ^η is differentiable wherever φ^Z is.

Assume that $\|\varphi^Z(t_0)\| > 0$ and t_0 is a regular point (see Subsection **6.1**). Let $N_0 = \sharp\{i :\ \varphi_i^Z(t_0) = 0\} < N$. Then, with a proper reordering, $\varphi_i^Z(t_0) = 0$, for

$i \in \{1, \ldots, N_0\}$ and $\varphi_i^Z(t_0) > 0$, for $i \in \{N_0 + 1, \ldots, N\}$. Since $\varphi_i^Z(t) \geq 0$ and t_0 is a regular point, from $\varphi_i^Z(t_0) = 0$ we get $(\varphi_i^Z)'(t_0) = 0$. We find the values of

$$(\varphi_i^Z)'(t_0) = -\mu + H_i(\varphi_i^\eta)'(t_0) + \sum_{j \neq i} w_j (\varphi_j^\eta)'(t_0). \tag{210}$$

We prove that $(\varphi_i^\eta)'(t_0) = 0$ for $i > N_0$ (if a potential is very far from the threshold then the neuron does not have a spike for a long time) and, therefore,

$$0 = -\mu + (H_i - w_i)(\varphi_i^\eta)'(t_0) + \sum_{j=1}^{N_0} w_j (\varphi_j^\eta)'(t_0), \quad i = 1, \ldots, N_0, \tag{211}$$

$$(\varphi_i^Z)'(t_0) = -\mu + \sum_{j=1}^{N_0} w_j (\varphi_j^\eta)'(t_0), \quad i = N_0 + 1, \ldots, N. \tag{212}$$

Let

$$h = \min_{N_0+1 \leq i \leq N} \varphi_i^Z(t_0). \tag{213}$$

We prove that for any $\Delta < h/(4\mu)$ and $i \in \{N_0 + 1, \ldots, N\}$ equality $\varphi_i^\eta(t_0 + \Delta) = \varphi_i^\eta(t_0)$ holds. Since $\widehat{Z}_i^{\mathbf{z}_n}(t_0) \Rightarrow \varphi_i^Z(t_0)$, we have $\widehat{Z}_i^{\mathbf{z}_n}(t_0) > h/2 > 2\mu\Delta$ a.s. for n large. We have

$$\begin{aligned}
&\mathbb{P}\{\eta_i^{\mathbf{z}_n}(\|\mathbf{z}_n\|(t_0 + \Delta)) > \eta_i^{\mathbf{z}_n}(\|\mathbf{z}_n\| t_0)\} \\
&\leq \mathbb{P}\{2\mu\Delta\|\mathbf{z}_n\| + \inf_{0 \leq s \leq \Delta}(X_i(\|\mathbf{z}_n\|(t_0 + s)) - X_i(\|\mathbf{z}_n\| t_0)) \leq 0\} \\
&\leq \mathbb{P}\{\mu\Delta\|\mathbf{z}_n\| + \inf_{0 \leq s \leq \Delta}(X_i^0(\|\mathbf{z}_n\|(t_0 + s)) - X_i^0(\|\mathbf{z}_n\| t_0)) \leq 0\} \\
&= \mathbb{P}\{\sup_{0 \leq s \leq \Delta} X_i^0(\|\mathbf{z}_n\| s) \geq \mu\Delta\|\mathbf{z}_n\|\}.
\end{aligned} \tag{214}$$

Thus, by Theorem 36.8 from Sato (1999), we have convergence

$$\eta_i^{\mathbf{z}_n}(\|\mathbf{z}_n\|(t_0 + \Delta)) - \eta_i^{\mathbf{z}_n}(\|\mathbf{z}_n\| t_0) \to 0 \tag{215}$$

in probability and convergence

$$\widehat{\eta}_i^{\mathbf{z}_n}(t_0 + \Delta) - \widehat{\eta}_i^{\mathbf{z}_n}(t_0) \to 0 \tag{216}$$

a.s., as $n \to \infty$. Thus, equality $\varphi_i^\eta(t_0 + \Delta) = \varphi_i^\eta(t_0)$ holds for $\Delta < h/(4\mu)$ and $(\varphi_i^\eta)'(t_0) = 0$.

Using Remark **5.7** we solve system (211) and get

$$(\varphi_i^\eta)'(t_0) = \frac{\mu}{H_i - w_i} \frac{1}{1 + \sum_{k=1}^{N_0} \frac{w_k}{H_k - w_k}}, \tag{217}$$

for $i = 1, \ldots, N_0$, and therefore,

$$(\varphi_i^Z)'(t_0) = -\mu + \frac{\mu}{1 + \sum_{k=1}^{N_0} \frac{w_k}{H_k - w_k}} \sum_{j=1}^{N_0} \frac{w_j}{H_j - w_j} = -\frac{\mu}{1 + \sum_{k=1}^{N_0} \frac{w_k}{H_k - w_k}}, \tag{218}$$

for $i = N_0 + 1, \ldots, N$.

Therefore, the process φ^Z is deterministic and piecewise linear. We have

$$\varphi_i^Z(0) \leq 1 \text{ and } (\varphi_i^Z)'(t_0) \leq -\frac{\mu}{1 + \sum_{k=1}^{N} \frac{w_k}{H_k - w_k}}, \quad i = 1, \ldots, N, \tag{219}$$

for any regular point t_0 such that $\varphi_i^Z(t_0) > 0$. Thus, from (177) we have that in time interval $(0, \mu^{-1}(1 + \sum_{k=1}^{N} \frac{w_k}{H_k - w_k}))$ process φ^Z reaches zero and stays there.

Let $\tau^{\mathbf{z}}(\varepsilon, B) = \inf\{t > \varepsilon : \ Z^{\mathbf{z}}(t) \in B\}$. Since fluid limits are stable, there exists $\kappa > 0$ such that for $V = \{\mathbf{z} \in \mathcal{Z} : \|\mathbf{z}\| < k_0\}$ we have

$$\sup_{\mathbf{z} \in V} \mathbb{E}\tau^{\mathbf{z}}(\varepsilon, V) < \infty. \tag{220}$$

§ 6.4. Comments on average drifts μ_i and signals b_{ij}

In this Subsection we discuss possible generalisations of conditions (162). As mentioned before, up until equation (208) the only condition we need on matrix B and vector $(\mu_i)_{i=1}^N$ is $b_{ij} > 0$ and $\mu_i > 0$ for $i, j \in \{1, \ldots, N\}$. However afterwards it is important for matrix B to be inversible. Additional conditions are also needed.

Let us start with examples. In a order to prove stability in the system of two inhibitory neurons it is sufficient to assume that the signals are smaller than the thresholds. However, in a system of three inhibitory neurons it is not sufficient. For instance, is $B = \begin{pmatrix} 8 & 2 & 6 \\ 2 & 8 & 6 \\ 6 & 6 & 8 \end{pmatrix}$ and drifts $\mu_i = -1$, $i = 1, 2, 3$, the first two neurons can form a stable system that "pushes" the potential of the third neuron to infinity. Thus, even though the matrix B is inversible and the signals are less than the thresholds, stability is not achieved.

Here is an example of sufficient conditions on matrix B and parameters μ_i to avoid such cases (we believe that they may be weakened).

- For every set $S \subseteq \{1, \ldots, N\}$ the matrix $B^S = (b_{ij})_{i,j \in S}$ is invertible and $a_i^S = \left(f^S (B^S)^{-1}\right)_i \geq 0$, $i \in S$, where $f^S = (\mu_i)_{i \in S}$;

- $\sum_{i\in S} a_i^S \sum_{j\notin S} b_{ij} < \sum_{j\notin S} \mu_j$.

We use notations and results of Subsection **6.3** to give an explanation for this conditions. Firstly, choose an arbitrary time $t_0 > 0$ and a set $S \subseteq \{1, \ldots, N\}$ such that $\varphi_i^Z(t_0) = 0$, for $i \in S$. Secondly, we try to solve an analogue of system of equation (211) to understand "how often neuron $i \in S$ spikes" in a current setup, i.e., find $(\varphi_i^\eta)'(t_0)$, for $i \in S$, which should be equal $\left(f^S(B^S)^{-1}\right)_i$.

The first condition is technical: if matrix B^S is degenerate then there might be no solutions to system $a^S B^S = f^S$, or the system may have an infinite number of solutions. If we assume that B is a strictly diagonally dominant matrix (i.e., $|b_{ii}| > \sum_{j\neq i} b_{ij}$), this particular problem is resolved. However, since $\eta_i^{\mathbf{z}}(t)$ is a non-decreasing function, value $(\varphi_i^\eta)'(t)$ must be non-negative for all $i \in \{1, \ldots, N\}$ and $t \geq 0$. Thus, we have the second part of the first condition.

The second condition has the following meaning: the subset of neurons with numbers $i \in S$ have spikes and we take the mean total value of signals sent to the rest of neurons, created by those spikes. We compare it with the mean total drift of neurons with numbers $j \notin S$. If the second condition holds then on average the subset of neurons with numbers $i \in S$ do not "push" the potentials of the rest of the neurons from the threshold and at least one neuron $j \notin S$ will eventually spike.

§ 6.5. Proof of existence of minorant measure

In this Subsetion we prove Lemma **5.9** by showing the existence of a lower bound for $\inf_{\mathbf{z}\in V} \mathbb{P}\{Z^{\mathbf{z}}(U) \in B\}$ where $V = \{\mathbf{z} \in \mathcal{Z} : \|\mathbf{z}\| < k_0\}$ (see the end of previous Subsection). By Theorem 19.2 from Sato (1999), the Lévy process $X(t)$ can be represented as a sum $X^1(t) + X^2(t)$ of two independent processes, a jump process $X^1(t)$ and a Gaussian process $X^2(t)$ with drift. We consider cases where at least one coordinate is close enough to zero. If all the coordinates of $\mathbf{z}$ are bounded away from zero then the proof follows similar lines.

Since random variables $\xi_{ij}^{(1)}$, $i, j \in [1, N]$, are strictly positive, there are constants $k_1^+, k_1^- > 0$ such that

$$p_1 = \mathbb{P}\{A_1\} \equiv \mathbb{P}\left\{(\xi_{ij}^{(1)})_{i,j=1}^N \in [k_1^-, k_1^+]^{N^2}\right\} > 0. \tag{221}$$

Without loss of generality, we assume that $z_1 < k_1^-/6$.

First, we bound the jump process $X^1(t)$ in the time interval $[0, 2]$, which includes the time interval $[0, U]$, and take time instant $t_0 \leq 1/2$ such that $\mathbb{P}\{X_1^1(t_0) < k_1^-/6\} > 0$. Denote

$$A_2 = \left\{ \max_{1 \leq i \leq N} (X_i^1(2)) < k_2 \right\} \cap \left\{ X_1^1(t_0) < k_1^-/6 \right\}, \tag{222}$$

and take a constant $k_2 > 0$ such that $p_2 = \mathbb{P}\{A_2\} > 0$. Next, we use the condition that the Gaussian process $X^2(t)$ is non-degenerate and none of its coordinates is a deterministic line. Thus, we denote

$$\begin{aligned} A_3 = &\left\{ \max_{1 \leq i \leq N} \sup_{0 \leq t \leq \frac{1}{2}} (X_i^2(t)) \leq k_3 - k_2 - k_1^- \right\} \cap \left\{ \min_{1 \leq i \leq N} \inf_{0 \leq t \leq \frac{1}{2}} (X_i^2(t)) \geq -\frac{k_1^-}{2} \right\} \\ &\cap \left\{ \inf_{0 \leq t \leq t_0} (X_1^2(t)) \leq -\frac{k_1^-}{3} \right\} \end{aligned} \tag{223}$$

and take a constant $k_3 > k_2 - k_1^-$ such that $p_3 = \mathbb{P}\{A_3\} > 0$. One can show that, given $A_2 \cap A_3$, the first spike occurs up to time t_0 and the second one can occur only after time $1/2$.

Denote the new set $D = A_1 \cap A_2 \cap A_3$. From independence of X^1, X^2 and $(\xi_{ij}^{(1)})_{i,j=1}^N$ we have $\mathbb{P}\{D\} = p_1 p_2 p_3 > 0$. We have

$$D \subseteq \left\{ \frac{k_1^-}{2} \leq Z_i^{\mathbf{z}}\left(\frac{1}{2}\right) + X_i^1(U) - X_i^1\left(\frac{1}{2}\right) \leq k_1^+ + k_2 + k_3,\ i = 1, \ldots, N \right\}. \tag{224}$$

We restrict ourselves to events without the second spike up to time U. Denote

$$G_2(t_1, t_2, k) = \left\{ \min_{1 \leq i \leq N} \inf_{t_1 \leq s \leq t_2} (X_i^2(s)) > -k \right\}. \tag{225}$$

Using (224), we get that, given $G_2\left(1/2, U, k_1^-/2\right) \cap D$, the second spike occurs after time U.

Let $K = k_1^+ + k_2 + k_3$. We prove that, for any point $\mathbf{y} \in (K, 2K)^N$ and a measurable set $\Delta \subset [0, K]^N$, there is a number $p > 0$ such that

$$\mathbb{P}\left\{ \{Z^{\mathbf{z}}(U) \in \mathbf{y} + \Delta\} \cap G_2\left(\frac{1}{2}, U, \frac{k_1^-}{2}\right) \cap D \right\} \geq p\lambda(\Delta), \tag{226}$$

where λ is the Lebesgue measure. Denote, $\widehat{y}_i = y_i - Z_i^{\mathbf{z}}(1/2) - (X_i^1(U) - X_i^1(1/2))$, for $i \in [1, N]$. Then we have

$$\{Z^{\mathbf{z}}(U) \in \mathbf{y} + \Delta\} = \left\{ X^2(U) - X^2\left(\frac{1}{2}\right) \in \widehat{\mathbf{y}} + \Delta \right\}. \tag{227}$$

Since X^2 is a Markov process, the events $\{X^2(U) - X^2(1/2) \in \widehat{\mathbf{y}} + \Delta\} \cap G_2\left(1/2, U, k_1^-/2\right)$ and D are independent, conditioned on a value of $\widehat{\mathbf{y}}$. Thus, we have

$$\begin{aligned}&\mathbb{P}\left\{\left\{X^2(U) - X^2\left(\frac{1}{2}\right) \in \widehat{\mathbf{y}} + \Delta\right\} \cap G_2\left(\frac{1}{2}, U, \frac{k_1^-}{2}\right) \cap D\right\}\\ &= \mathbb{E}\left(\mathbb{P}\left\{\left\{X^2(U) - X^2\left(\frac{1}{2}\right) \in \widehat{\mathbf{y}} + \Delta\right\} \cap G_2\left(\frac{1}{2}, U, \frac{k_1^-}{2}\right) \mid \widehat{\mathbf{y}}\right\} \mathbb{P}\{D \mid \widehat{\mathbf{y}}\}\right)\end{aligned} \tag{228}$$

Next, we need a technical lemma regarding a monotonicity property of the Brownian bridge.

Lemma 6.4. *For any $t, k > 0$ and $\Delta \subset [0,\infty)^N$ we have*

$$\mathbb{P}\{G_2(0,t,k) \mid \ X^2(t) \in \Delta\} \geq \mathbb{P}\{G_2(0,t,k) \mid \ X^2(t) = \boldsymbol{0}\} > 0. \tag{229}$$

Given D, we have $\widehat{y}_i \geq 0$, for $i \in [1, N]$, and we can use Lemma **6.4** to obtain

$$\begin{aligned}&\mathbb{P}\left\{\left\{X^2(U) - X^2\left(\frac{1}{2}\right) \in \widehat{\mathbf{y}} + \Delta\right\} \cap G_2\left(\frac{1}{2}, U, \frac{k_1^-}{2}\right) \mid \widehat{\mathbf{y}}\right\}\\ &\overset{a.s.}{\geq} \mathbb{P}\left\{X^2(U) - X^2\left(\frac{1}{2}\right) \in \widehat{\mathbf{y}} + \Delta \mid \widehat{\mathbf{y}}\right\} \times\\ &\times \mathbb{P}\left\{G_2\left(\frac{1}{2}, U, \frac{k_1^-}{2}\right) \mid \ X^2(U) - X^2\left(\frac{1}{2}\right) = \mathbf{0}\right\}.\end{aligned} \tag{230}$$

The density of $X^2(t)$ is bounded away from zero on any compact set, and X^2 and U are independent. Therefore, there exists $p_4 > 0$ such that, given D, for a measurable set $\Delta \subseteq [0, K]^N$ we have

$$\mathbb{P}\left\{X^2(U) - X^2\left(\frac{1}{2}\right) \in \widehat{\mathbf{y}} + \Delta \mid \widehat{\mathbf{y}}\right\} \overset{a.s.}{\geq} p_4\lambda(\Delta). \tag{231}$$

Denote $p_4' = p_4\mathbb{P}\left\{G_2\left(\frac{1}{2}, U, \frac{k_1^-}{2}\right) \mid \ X^2(U) - X^2\left(\frac{1}{2}\right) = \mathbf{0}\right\} > 0$. Combining altogether, we get that if $z_1 < k_1^-/6$ then

$$\mathbb{P}\{Z^{\mathbf{z}}(U) \in \mathbf{y} + \Delta\} \geq p_1p_2p_3p_4'\lambda(\Delta), \tag{232}$$

for $\mathbf{y} \in (K, 2K]$ and $\Delta \subseteq [0, K]^N$. Therefore, we proved the existence of a minorant measure, and thus, proved Lemma **5.9**. Together with Lemma **5.8**, we have the proof of Harris positive recurrence of process $(Z_1^z(t), \ldots, Z_N^z(t))$ and, therefore, the proof of Theorem **5.6**.

SECTION 7

Spike-analysis in two particular cases

In this Section we consider two very particular and simple examples of two-dimensional neuron networks ($N = 2$) and analyse the time-instants when spikes occur. In a more general model we proved Harris positive recurrence. However, we did not obtain results on the behaviour of the model in the stationary regime. Firstly, we dedicate this Section to illustrate what kind of explicit formulae can be obtained in these models. Similar results in continuous-time setting can be found in Cottrell (1992). Secondly, we use this Section as a setup for a future research of sensitivity of stationary regime to the distributions of the input noise and the signals. Namely, in our general model the results depend mostly on the mean values H_i, w_i and μ. Thus, we do not see any dependence on the tail-distribution of the signals. Should something change if $\xi_{ij}^{(n)}$ has a heavy-tail distribution? The final statement of this Section (Remark **7.4**) we illustrate our first results in finding such sensitivity.

§ 7.1. Constant input

In this Subsection we assume that the process $X(t)$, which represents combined internal and external noise, is deterministic and

$$X_i(t) = -t, \text{ for } t \geq 0 \text{ and } i = 1, 2. \tag{233}$$

For the distribution of signals between neurons we assume

$$\xi_{ij}^{(1)} \in \mathbb{Z}_+ \text{ and } \xi_{ii}^{(1)} = 1, \text{ for } i, j = 1, 2, \text{ and } \; w_i = \mathbb{E}\xi_{ij}^{(1)} < 1 \text{ for } i \neq j. \tag{234}$$

We assume here that the travel time of signals between neurons is zero. As was mention in the introduction of Part **II**, it is important to deal with uncertainty in the order of spikes. Let $Z_1(0) = Z_2(0) = 0$. We assume that each time $t \geq 0$, when $Z_1(t) = Z_2(t) = 0$, the neuron which spikes first is chosen with equal probabilities

1/2 and independently of everything else. Further, if neuron 1 is the spiking neuron and the current signal $\xi_{12}^{(n)}$ from 1 to 2 equals zero then we assume that neuron 2 has a spike directly after neuron 1 spiked.

In this case we want to analyse the time-instants when a neuron 'has a potential to have a spike'. Define independent sequence $\{\zeta_n\}_{n=1}^{\infty}$ of i.i.d. r.v.'s with Bernoulli distribution with parameter 1/2. Let $\tau_1(0) = \tau_2(0) = 0$. For $n \geq 1$ we have

$$\tau_1(n) = \tau_1(n-1) + \begin{cases} 1, & \text{if } \tau_1(n-1) < \tau_2(n-1), \\ \zeta_n + (1-\zeta_n)\xi_{21}^{(n)}, & \text{if } \tau_1(n-1) = \tau_2(n-1), \\ \xi_{21}^{(n)}, & \text{if } \tau_1(n-1) > \tau_2(n-1), \end{cases} \tag{235}$$

$$\tau_2(n) = \tau_2(n-1) + \begin{cases} 1, & \text{if } \tau_2(n-1) < \tau_1(n-1) \\ 1 - \zeta_n + \zeta_n\xi_{12}^{(n)}, & \text{if } \tau_2(n-1) = \tau_1(n-1), \\ \xi_{12}^{(n)}, & \text{if } \tau_2(n-1) = \tau_1(n-1). \end{cases} \tag{236}$$

We call time-instants $\{\tau_1(n)\}_{n=1}^{\infty}$ and $\{\tau_2(n)\}_{n=1}^{\infty}$ the potential times for spikes of neuron 1 and 2, respectively. Indeed, if after time $\tau_1(n-1)$ neuron 2 would stop sending signals, neuron 1 will have a spike at time $\tau_1(n)$.

Lemma 7.1. *We have*

$$\frac{\tau_1(n)}{n} \to \frac{1 - w_1 w_2}{2 - w_1 - w_2} \leftarrow \frac{\tau_2(n)}{n} \quad a.s. \text{ as } n \to \infty. \tag{237}$$

Proof. Let $\nu = \inf\{n > 0 : \ \tau_1(n) = \tau_2(n)\}$. We can represent the timeline as a union of i.i.d. cycles that start and end at instants $\tau_1(n) = \tau_2(n)$. First assume that $\tau_1(1) < \tau_2(1)$. Then, for $n < \nu$, we have that $\tau_1(n+1) = \tau_1(n) + 1$ and $\tau_2(n+1) = \tau_2(n) + \xi_{12}^{(n+1)}$. Thus, sequence $\{\tau_2(n+1) - \tau_1(n+1)\}_{n=0}^{\nu-1}$ is a left-continuous random walk (i.e. a random walk such that its negative increments can only be equal to -1) with a negative drift $(w_1 - 1)$ and this random walk is killed at the moment of reaching zero. It is straightforward to calculate that

$$\mathbb{E}\left(\nu - 1 \mid \tau_2(1) - \tau_1(1) = k\right) = \frac{k}{1 - w_1} \text{ for } k \geq 1. \tag{238}$$

A similar argument applies to the case where $\tau_1(1) > \tau_2(1)$. Let $\chi \in \{1, 2\}$ be the number of the neuron which has a spike at time 0. We have

$$\mathbb{E}(\nu \mid \chi = 1, \ \xi_{12}^{(1)} = 1) = \mathbb{E}(\nu \mid \chi = 2, \ \xi_{21}^{(1)} = 1) = 1, \tag{239}$$

$$\mathbb{E}(\nu|\ \chi = 1,\ \xi_{12}^{(1)} = 0) = 1 + \frac{1}{1 - w_2}, \tag{240}$$

$$\mathbb{E}(\nu|\ \chi = 2,\ \xi_{21}^{(1)} = 0) = 1 + \frac{1}{1 - w_1}, \tag{241}$$

$$\mathbb{E}(\nu|\ \chi = 1,\ \xi_{12}^{(1)} = k > 1) = 1 + \frac{k-1}{1 - w_1}, \tag{242}$$

$$\mathbb{E}(\nu|\ \chi = 2,\ \xi_{12}^{(1)} = k > 1) = 1 + \frac{k-1}{1 - w_2}. \tag{243}$$

Let $p_{1k} = \mathbb{P}\{\xi_{12}^{(1)} = k\}$ and $p_{2k} = \mathbb{P}\{\xi_{21}^{(1)} = k\}$, for $k \in \mathbb{Z}^+$. Then we have

$$\mathbb{E}\nu = 1 + \frac{1}{2}\left(\frac{p_{10} + \sum_{k=2}^{\infty}(k-1)p_{2k}}{1 - w_2} + \frac{p_{20} + \sum_{k=2}^{\infty}(k-1)p_{1k}}{1 - w_1}\right). \tag{244}$$

Since $\sum_{k=2}^{\infty}(k-1)p_{ik} = w_i - 1 + p_{i0} = p_{i0} - (1 - w_i)$, we have

$$\mathbb{E}\nu = 1 + \frac{1}{2}\left(\frac{p_{10} + p_{20}}{1 - w_1} + \frac{p_{10} + p_{20}}{1 - w_1} - 2\right) = \frac{p_{10} + p_{20}}{2}\left(\frac{1}{1 - w_1} + \frac{1}{1 - w_2}\right). \tag{245}$$

Now, if $\tau_{3-\chi}(1) = 0$ (the first signal equals zero) then $\tau_1(\nu) = \nu - 1$. Otherwise, we have $\tau_1(\nu) = \nu$. Thus, we have

$$\mathbb{E}\tau_1(\nu) = \frac{p_{10} + p_{20}}{2}\left(\frac{1}{1 - w_1} + \frac{1}{1 - w_2}\right) - \frac{p_{10} + p_{20}}{2}. \tag{246}$$

Finally, we have

$$\frac{\tau_1(n)}{n} \xrightarrow{a.s.} \frac{\mathbb{E}\tau_1(\nu)}{\mathbb{E}\nu} = \frac{\frac{1}{1-w_1} + \frac{1}{1-w_2} - 1}{\frac{1}{1-w_1} + \frac{1}{1-w_2}} = \frac{1 - w_1 w_2}{2 - w_1 - w_2},\ \text{as } n \to \infty. \tag{247}$$

□

Now, time-instant $\min(\tau_1(n), \tau_2(n))$ corresponds to the time when the neuron network has a spike number $(n+1)$. From Lemma **7.1**, we have

$$\frac{\min(\tau_1(n), \tau_2(n))}{n} \to \frac{1 - w_1 w_2}{2 - w_1 - w_2}\ \text{a.s. as } n \to \infty. \tag{248}$$

Thus, we can say that in the stationary regime the mean-time between any two subsequent spikes is $(1 - w_1 w_2)/(2 - w_1 - w_2)$. As for the individual neurons, let us analyse $\eta_1(\nu)$, the number of actual spikes of neuron 1 at time-instants $\{\tau_1(n)\}_{n=0}^{\nu-1}$.

$$\mathbb{E}(\eta_1(\nu)|\ \chi = 1,\ \xi_{12}^{(1)} = 1) = \mathbb{E}(\eta_1(\nu)|\ \chi = 1,\ \xi_{12}^{(1)} = 0) = 1, \tag{249}$$

$$\mathbb{E}(\eta_1(\nu)|\ \chi = 2,\ \xi_{21}^{(1)} = 0) = \frac{1}{1-w_1}, \tag{250}$$

$$\mathbb{E}(\eta_1(\nu)|\ \chi = 1,\ \xi_{12}^{(1)} = k > 1) = 1 + \frac{k-1}{1-w_1}, \tag{251}$$

$$\mathbb{E}(\eta_1(\nu)|\ \chi = 2,\ \xi_{21}^{(1)} = 1) = \mathbb{E}(\eta_1(\nu)|\ \chi = 2,\ \xi_{12}^{(1)} = k > 1) = 0. \tag{252}$$

Thus, we have

$$\begin{aligned}\mathbb{E}\eta_1(\nu) &= \frac{1}{2}\left(1 + \frac{\sum_{k=2}^{\infty}(k-1)p_{1k}}{1-w_1}\right) + \frac{p_{20}}{2(1-w_1)} \\ &= \frac{1}{2}\left(1 + \frac{w_1 - 1 + p_{10} + p_{20}}{1-w_1}\right) = \frac{p_{10}+p_{20}}{2(1-w_1)}.\end{aligned} \tag{253}$$

Therefore, the mean-number of spikes neuron 1 has before time $\tau_1(\nu)$ equals $(p_{10}+p_{20})/(2(1-w_1))$. Finally, for the number of spikes $N_1(t)$ neuron 1 has before time t we get

$$\frac{N_1(t)}{t} \to \frac{\mathbb{E}\eta_1(\nu)}{\mathbb{E}\tau_1(\nu)} = \frac{1-w_2}{1-w_1w_2}, \text{ a.s. as } t \to \infty. \tag{254}$$

Thus, we can say that in the stationary regime the mean-time between any two subsequent spikes of neuron 1 is $(1-w_1w_2)/(1-w_2)$.

Remark 7.2. There is another way to resolve uncertainty in the order of spikes when both membrane potentials reach zero. We again assume that each time $t \geq 0$, when $Z_1(t) = Z_2(t) = 0$, the neuron which spikes first is chosen with equal probabilities 1/2 and independently of everything else. Further, if neuron 1 is the spiking neuron and the current signal $\xi_{12}^{(n)}$ from 1 to 2 equals zero then we assume that both potentials change value to 1 and proceed until the next spike. Thus, we assume that, if neuron 2 has a spike directly after neuron 1 spiked, signal $\xi_{21}^{(n)}$ does not reach neuron 1. Thus, the dynamics change to

$$\tau_1(n) = \tau_1(n-1) + \begin{cases} 1, & \text{if } \tau_1(n-1) < \tau_2(n-1), \\ \zeta_n + (1-\zeta_n)\min(1, \xi_{21}^{(n)}), & \text{if } \tau_1(n-1) = \tau_2(n-1), \\ \xi_{21}^{(n)}, & \text{if } \tau_1(n-1) > \tau_2(n-1), \end{cases} \tag{255}$$

$$\tau_2(n) = \tau_2(n-1) + \begin{cases} 1, & \text{if } \tau_2(n-1) < \tau_1(n-1) \\ 1 - \zeta_n + \zeta_n \min(1, \xi_{12}^{(n)}), & \text{if } \tau_2(n-1) = \tau_1(n-1), \\ \xi_{12}^{(n)}, & \text{if } \tau_2(n-1) = \tau_1(n-1). \end{cases} \tag{256}$$

§ 7.2. Poisson input

Let $N_1(t)$ and $N_2(t)$ be two independent right-continuous Poisson processes with intensities λ_1 and λ_2, respectively. In this Subsection we assume that

$$X_i(t) = -N_i(t), \text{ for } i = 1, 2. \tag{257}$$

For the distribution of signals between neurons we assume

$$\xi_{ij}^{(1)} \in \mathbb{Z}_+ \text{ and } \xi_{ii}^{(1)} = 1, \text{ for } i, j = 1, 2, \tag{258}$$

and

$$w_1 = \mathbb{E}\xi_{12}^{(1)} < \frac{\lambda_2}{\lambda_1} \text{ and } w_2 = \mathbb{E}\xi_{21}^{(1)} < \frac{\lambda_1}{\lambda_2}. \tag{259}$$

Under these assumptions we have the following result.

Proposition 7.3. *Assume* (257) - (259) *hold. For any fixed initial value* $Z(0)$*, process* $Z(t)$ *is positive recurrent with an atom at* $(1, 1)$*.*

Assume $Z(0) = (z_1, z_2)$, where z_i, $i = 1, 2$ are positive integer. Then for any spiking time T_{ik} we have $Z_i(T_{ik}) = 0$ and $Z_i(T_{ik} - 0) = Z_i(T_{ik} + 0) = 1$, with probability 1. For the purposes of this Subsection, we define a left-continuous process $Z'(t)$ which, apart from spiking times, has the same properties as $Z(t)$. However, for any time $T'_{ik} \stackrel{d}{=} T_{ik}$ we have $Z'_i(T'_{ik} - 0) = Z'_i(T'_{ik}) = Z'_i(T'_{ik} + 0) = 1$. Therefore, if process $Z'(t)$ starts at positive integer values, set

$$Q = \{(z_1, z_2) \in \mathbb{Z}^2 : \ \min(z_1, z_2) = 1\} \tag{260}$$

is a reflective barrier. Further, set Q is absorbing: once the process hits this set it will remain there afterwards. Thus, our 2-dimensional jump process may be viewed as 1-dimensional.

Let $N(t)$ be a Poisson process with intensity $\lambda_1 + \lambda_2$ with arrival times $\{T_k\}_{k=1}^{\infty}$ independent of $\{\{\xi_{i,j}^{(k)}\}_{k=1}^{\infty}\}_{i,j=1}^{2}$. Let $\{\chi_k\}_{k=1}^{\infty}$ be a independent sequence of i.i.d. r.v.'s. We call χ_k a mark for arrival time T_k. We assume that

$$\mathbb{P}\{\chi_k = i\} = \frac{\lambda_i}{\lambda_1 + \lambda_2}, \text{ for } i = 1, 2, \ k \geq 1, \tag{261}$$

and if $\chi_k = i$ we call time T_k a 'i-point'. Now $(Z'_1(t), Z'_2(t))$ is a 2-dimensional Markov jump process that has jumps at points T_k of $N(t)$. Initially, $Z'_1(0) = z_1 \geq 1$

and $Z_2'(0) = z_2 \geq 1$ are positive integers and, for any $t > 0$, both components take only such values. If, for $k \geq 1$, time T_k is a 1-point ($\chi_k = 1$), we let

$$Z_1'(T_k + 0) = \max(Z_1'(T_k) - 1, 1) \text{ and } Z_2'(T_k + 0) = Z_2'(T_k) + \xi_{12}^{(k)} I[Z_1'(T_k) = 1]. \tag{262}$$

By the symmetry, if, for $k \geq 1$, time T_k is a 2-point ($\chi_k = 2$), we let

$$Z_2'(T_k + 0) = \max(Z_2'(T_k) - 1, 1) \text{ and } Z_1'(T_k + 0) = Z_1'(T_k) + \xi_{21}^{(k)} I[Z_2'(T_k) = 1]. \tag{263}$$

These formulae have a natural meaning: say, if T_k is a 1-point, then the first component of $Z'(t)$ decreases by 1 if it does not touch 0; if it touches zero, neuron 1 spikes and its potential increases by $\xi_{11}^{(k)} = 1$.

Let

$$\nu_0 = 0 \text{ and } \nu_m = \inf\{k > \nu_{m-1} : \ Z_1'(T_k + 0) = Z_2'(T_k + 0) = 1\}, \text{ for } m \geq 1. \tag{264}$$

Then the random vectors $\langle \nu_{m+1} - \nu_m, \ \{Z'(T_{n+\nu_m}), 1 \leq n \leq \nu_{m+1} - \nu_m\}\rangle$ are mutually independent, for $m \geq 0$, and identically distributed. We find now the mean length of a typical cycle $\nu_{m+1} - \nu_m$, $m \geq 1$. Given that, we find further the mean number of 1-spikes and 2-spikes within a cycle. Based on that, we find the limiting average number of 1- and 2-spikes.

For convenience, assume $Z_1'(0) = Z_2'(0) = 1$. Then we have

$$\mathbb{E}(\nu_1 | \ \chi_1 = 1, \ \xi_{12}^{(1)} = 0) = \mathbb{E}(\nu_1 | \ \chi_1 = 2, \ \xi_{21}^{(1)} = 0) = 1. \tag{265}$$

Assume $\chi_1 = 1$ and $\xi_{12}^{(1)} = n \geq 1$. Then, for $k \in \{1, \dots, \nu_1 - 1\}$, we have $Z_1'(T_k + 0) = 1$ and

$$Z_2'(T_{k+1} + 0) - Z_2'(T_k + 0) = \xi_{12}^{(k+1)} I[\chi_{k+1} = 1] - I[\chi_{k+1} = 2]. \tag{266}$$

From (259) we have

$$\mathbb{E}\left(\xi_{12}^{(k+1)} I[\chi_{k+1} = 1] - I[\chi_{k+1} = 2]\right) = \frac{w_1\lambda_1 - \lambda_2}{\lambda_1 + \lambda_2} < 0. \tag{267}$$

Thus, sequence $\{Z_2'(T_k + 0) - 1\}_{k=1}^{\nu_1}$ is a left-continuous random walk (i.e. a random walk such that its negative increments can only be equal to -1) with a negative drift $(w_1\lambda_1 - \lambda_2)/(\lambda_1 + \lambda_2)$ and this random walk is killed at the moment of reaching zero. It is straightforward to calculate that

$$\mathbb{E}\left(\nu_1 - 1 |\ \chi_1 = 1,\ \xi_{12}^{(1)} = n\right) = \frac{n}{\lambda_2 - w_1\lambda_1}(\lambda_1 + \lambda_2) \text{ for } k \geq 1. \tag{268}$$

A similar argument applies to the case $\chi_1 = 2$ and $\xi_{21}^{(1)} = n \geq 1$. Thus, we get

$$\mathbb{E}\nu_1 = 1 + \frac{w_1\lambda_1}{\lambda_2 - w_1\lambda_1} + \frac{w_2\lambda_2}{\lambda_1 - w_2\lambda_2}. \tag{269}$$

As for individual neurons, denote $\eta_i(\nu_1)$ as the number of spikes neuron i has in time interval $[0, T_{\nu_1}]$. If $\chi_1 = 1$ the $\eta_1(\nu_1) = \nu_1$ and $\eta_2(\nu_1) = 0$ (similar for $\chi_1 = 2$). Therefore, we get

$$\mathbb{E}\eta_1(\nu_1) = \mathbb{E}(\nu_1;\ \chi_1 = 1) = \frac{\lambda_1}{\lambda_1 + \lambda_2} + \frac{w_1\lambda_1}{\lambda_2 - w_1\lambda_1} \tag{270}$$

and

$$\mathbb{E}\eta_2(\nu_1) = \frac{\lambda_2}{\lambda_1 + \lambda_2} + \frac{w_2\lambda_2}{\lambda_1 - w_2\lambda_2}. \tag{271}$$

Finally, in the stationary regime the average number of i-spikes in a unit of time is $\mathbb{E}\eta_1(\nu_1)/\mathbb{E}T_{\nu_1} = \mathbb{E}\eta_1(\nu_1)/(\mathbb{E}\nu_1\mathbb{E}T_1)$, and the average time between i-spikes is the inverse.

Remark 7.4. In the special cases studied in this Section the average time between i-spikes in the stationary regime depends only on the first moments of random values that govern our models (e.g., w_i and λ_i). However, if, for example, we change condition (258) by letting $\mathbb{P}\{\xi_{11}^{(1)} = 2\} > 0$, the limiting result becomes more sensitive to the distribution of $\xi_{12}^{(1)}$. Also, set $Q = \{(z_1, z_2) \in \mathbb{Z}^2 :\ \min(z_1, z_2) = 1\}$ stops being absorbing and the jump process becomes essentially 2-dimensional. Here is a brief description of our results in this case. Let $r, p = 1 - q, p_0 = 1 - q_0 \in (0, 1)$ and assume that

$$\mathbb{P}\{\xi_1^{(11)} = 2\} = r = 1 - \mathbb{P}\{\xi_1^{(11)} = 1\} \in (0, 1), \tag{272}$$

$$\mathbb{P}\{\xi_1^{(12)} = 0\} = p_0 = 1 - q_0 \text{ and } \mathbb{P}\{\xi_1^{(12)} = k |\ \xi_1^{(12)} \geq 1\} = pq^{k-1},\ \text{for } k \geq 1. \tag{273}$$

Denote $\alpha_i = \lambda_i/(\lambda_1 + \lambda_2)$. Assume stability conditions

$$\frac{1 + r}{\alpha_1} > \frac{q_0}{p\alpha_2} \text{ and } \frac{1}{\alpha_1} > \frac{w_2}{\alpha_2}. \tag{274}$$

Through analysis of our process at the embedded epochs T_k such that $Z_1(T_k) = 1$, we find

$$\mathbb{E}\eta_1(\nu_1) = \frac{\alpha_1}{1 - \widehat{p}}, \tag{275}$$

where

$$\widehat{p} = \frac{1 - qx}{(1 - q)x}, \tag{276}$$

$$x = \alpha_2 \frac{1 + r + (1 - r)q_0 + (q - (1 - r)q_0 - r)\alpha_2 + \sqrt{D}}{2(q_0 - (1 + r)(q_0 - q)\alpha_2 + r(q_0 - q)\alpha_2^2)}, \tag{277}$$

$$\begin{aligned} D = {} & (1 + r + (1 - r)q_0 + (q - (1 - r)q_0 - r)\alpha_2)^2 \\ & - 4(q_0 - (1 + r)(q_0 - q)\alpha_2 + r(q_0 - q)\alpha_2^2). \end{aligned} \tag{278}$$

Additionally, we find $\mathbb{E}\eta_2(\nu_1)$. However, we did not obtain $\mathbb{E}\nu_1$ and the limiting average of time between i-spikes. Nonetheless, the ratio $\mathbb{E}\eta_2(\nu_1)/\mathbb{E}\eta_1(\nu_1)$ contains the information about the proportion between the number of 1- and 2-spikes and shows certain sensitivity with respect to distribution of $\xi_{12}^{(1)}$.

SECTION 8

Future research

In this Section we discuss possible directions for future research based on our results. For the Cat-and-Mouse Markov chain, in order to achieve a new asymptotic behaviour of the mouse we allow the increments $\xi_n^{(2)}$ to potentially have an infinite second moment and in return we restrict the tail-distribution to be regularly varying. However, in order to acquire our result, we restrict the increments of the cat $\xi_n^{(1)}$ to take values ± 1 with equal probabilities $1/2$. The main reason for this restriction is the possibility to represent $\tau_k^{(1)}$, hitting time of a certain level $k \in \mathbb{Z}^+$, as a sum of k independent identically distributed random variables. It is our belief that this model can be generalised to the case where $\xi_n^{(1)}$ has aperiodic distribution with zero mean and finite variance. Nevertheless, the analysis of characteristic functions suggests that, in generalisation of the distribution of $\xi_n^{(1)}$, the finiteness of $\mathbf{Var}\xi_n^{(2)}$ plays an important role. More precisely, in the approximation of characteristic function of $\tau_m^{(1)}$ we get a summand of order m^2. Thus, to get rid of this summand we need series $\sum \Big(m^2 \mathbb{P}\{\xi_1^{(2)} = m\} \Big)$ to converge. As a risky plan, we could consider the increments of the mouse with an infinite first moment (and regularly varying tails), although right now we lack the intuition about the possible behaviour. Generalising the increments of the cat beyond finiteness of the second moment seems the hardest direction, since we loose instruments regarding times between jumps of the mouse.

As for the other directions for the research on the Cat-and-Mouse Markov chain, in our work we only considered the case when the "agents" are processes on $\mathbb{Z}$ in discrete time. Litvak and Robert (2012) also considered the cases when "agents live" on $\mathbb{Z}^2$ and $\mathbb{Z}^+$, as well as the cases of continuous-time Markov chains. We believe that the instruments developed by Uchiyama (2011a) can help us to produce a generalisation in a case of $\mathbb{Z}^2$. More precisely, we want to prove that if $\xi_n^{(1)}$ has aperiodic distribution with zero mean and finite variance and $\mathbf{Var}\xi_n^{(2)}$ then there is a process $Z(t)$ such that

$$\left\{\frac{M([\exp(nt)]}{n^{1/2}},\ t \geq 0\right\} \overset{D}{\Rightarrow} \{Z(t)\ t \geq 0\}, \quad \text{as } n \to \infty. \tag{279}$$

It would be very interesting to acquire generalised results in aforementioned areas as well. As a risky plan, one might consider agents, as an objects with positive volume, and interactions caused by non-empty intersections.

So far we were discussing only the results on a limiting behaviour of the mouse. Meanwhile, it would be very interesting to acquire results on joint convergence. For example, in the "standard setting", we would like to verify if there is a Wiener process $B(t)$ such that

$$\left\{\left(\frac{C([nt])}{n^{1/2}}, \frac{M([nt])}{n^{1/4}}\right),\ t \geq 0\right\} \overset{D}{\Rightarrow} \left\{(B(t), A^{(2)}(E^{(2)}(t)))\ t \geq 0\right\}, \quad \text{as } n \to \infty. \tag{280}$$

As a natural generalisation to the Dog-and-Cat-and-Mouse Markov chain we would like to consider the cases where the dog and the mouse have general distribution with finite second moment. The most challenging in this model is to generalise the distribution of the cat. Our analysis of trajectories was mainly based around tracking every moment, when the cat and the mouse can meet, and constructing a random walk which approximate the limiting behaviour of the mouse. The general distribution of the cat introduces a complex dependence between the current position and the next meeting time-instant of the cat and the mouse.

For a hierarchy chain of an arbitrary length we would like to strengthen our result to weak convergence of a processes. However, we do not believe that it should be done the same way it was achieved for the Dog-and-Cat-and-Mouse Markov chain. When the number of agents becomes four we already see a great increase in the amount of combinatorial computations. Thus, without a proper induction argument our technique seems ineffective.

As a separate direction, we would like to consider non-linear hierarchy structures. First, the agents can form tree-like system of relationships where the branches are independent conditioned on the root. As a simple example, it is interesting to consider a set of one cat and two mice and acquire a result in the form

$$\left\{\left(\frac{M_1([nt])}{n^{1/4}}, \frac{M_2([nt])}{n^{1/4}}\right),\ t \geq 0\right\} \overset{D}{\Rightarrow} \left\{(A_1^{(2)}(E_1^{(2)}(t)), A_2^{(2)}(E_2^{(2)}(t)))\ t \geq 0\right\}, \tag{281}$$

as $n \to \infty$. Second, the system of relationships can be 'inverse' to the previous case and the branches are independent if they do not intersect. Even in the case of two cats and one mouse it is not obvious if the mouse has the same scaling $n^{1/4}$, although it is clear that the number of jumps for the mouse will only increase in comparison with the standard case. It is important to notice that the complex dependence between the current position of the system and the next time-instant, when the mouse will meet one of the cats, forces us to find a new method of analysis with regards to trajectories of the system. Finally, the model can be modified into a fork-join setting, however, this is an extremely challenging case.

So far we were dealing only with the cases where the jumps of all agents are mutually independent random variables. One of the earliest works on Cat-and-Mouse Markov chain by Coppersmith *et al.* (1993) is motivated by on-line algorithms and deals with a case where there is a clear competition between the cat and the mouse. Thus, it would be very intriguing to introduce a various dependencies between the jumps of an agent, history of meeting time-instants and current position (for example, the mouse can try to find places which the cat rarely visits). A relatively simple modification would be to forbid agents X_i and X_{i+1} to jump at the same place after a meeting time-instant, which would prevent us from using the same construction in the Dog-and-Cat-and-Mouse case, since all three agents will not meet together. A combination of similar modifications and a setting of one cat and large amount of mice can lead to an interesting problem of dispersing particles on a line.

For the neuron networks we considered the case where the membrane potentials behave like spectrally positive (the increments at discontinuity points are positive) Lévy processes $X_i(t)$ when they are away from zero. When reaching zero, neurons send positive signals $\xi_{ij}^{(k)}$ to other neurons (all neurons are inhibitory). It is important to mention that our conditions on the average signals b_{ij} and the drifts $\mu_i = -\mathbb{E}X_i(1)$ of Lévy processes are far from necessary for stationarity and it is very intriguing to explore further cases and find necessary and sufficient conditions. One the most interesting moments here is that the fluid limit may become 'essentially random'. In the proof of positive recurrence, for each neuron i we saw: if $\varphi_i^Z(t_0) = 0$ (membrane potential is relatively close to its threshold) then the rate of spikes $(\varphi_i^\eta)'(t_0)$ is derived from a system of linear equations

$$(\varphi^{\eta,S})'(t_0)B^S = \mu^S, \text{where } S = \{i: \ \varphi_i^Z(t_0) = 0\}, \tag{282}$$

$$\varphi^{\eta,S} = (\varphi_i^\eta: \ i \in S), \ B^S = (b_{kj})_{k,j\in S}, \text{ and } \mu^S = (\mu_i: \ i \in S). \tag{283}$$

If the matrix B^S is non-invertible then, depending of μ^S, there might be an uncountably infinite amount of solutions which will lead to a situation where the influence on φ_j^Z, $j \notin S$, might be random with unclear distribution. Our current conditions are simply an example which insures, that the matrices B^S are invertible, and provides a solution in a clear and simple form.

When $|S| \leq 2$, the necessary and sufficient conditions for the existence of an inverse $(B^S)^{-1}$ are clear. Let us restrict matrix B to these conditions. Then, for the time-region where $|S| \leq 2$, the fluid limits $\varphi^Z(t)$ and $\varphi^\eta(t)$ are random, and yet they are piecewise deterministic and linear. However, when set S contains three or more elements we may get an uncountable branching for the trajectories of fluid limits (such phenomenon is sometimes called *scattering*). When membrane potentials are near their thresholds, fluid limit condense the trajectories into the zero-line. Thus, a deterministic fluid limit sometimes cannot capture the intricate relationships of trajectories for neurons with number from set S. Similar effects were studied by Foss and Kovalevskii (1999) with discrete branching and Kovalevskii *et al.* (2005) with uncountable branching.

The next step of our research is to consider networks of excitatory neurons, meaning that the signals between neurons are negative. In the term of proof of positive recurrence, this model is easier. However, this introduces the situations where multiple spikes occur at the same time. There are different ways to model the system to resolve such events. It is not clear how this will influence the spiking rate for each neuron in a stationary regime. Nevertheless, if all signals are non-zero, then we have a non-zero chance for the whole system to spike at the same time. This introduces regeneration events and gives us a basis for a very interesting analysis. From the technical standpoint, if in the inhibitory case we could simply write the influence of spikes in the form $\sum_{j=1}^{N}\sum_{k=1}^{\eta_j^{\mathbf{z}}(t)} \xi_{ji}^{(k)}$, with the introduction of excitatory neurons this influence becomes strongly state-dependent (instead of $\xi_{ji}^{(k)}$ there will be a possibly greater random variable which depends on the values of every other current spike).

Our main goal is to move to networks with certain ratio between numbers of inhibitory and excitatory neurons and acquire mean-field results. Let assume that the number of inhibitory neurons is N^+ and the number of excitatory neurons is N^-. For $N = N^+ + N^-$, we assume that the average signals $b_{ij}^{(N)}$ are of order N^{-1}, as $N \to \infty$. In a case, where N^+/N^- converges to a certain $\alpha > 0$, as $N \to \infty$, we want to find a distribution of $\lim_{N\to\infty} Z_i^{(N)}(t)$. A biggest challenge in the presence of both types of neurons is that we loose any kind of monotonicity (in the inhibitory case we at least have a partial monotonicity). Additionally, the aforementioned problem of simultaneous spikes becomes much more prevalent and resolutions are less intuitive.

Another important direction of our research is the acquisition of precise characteristics of the networks in the stationary regime. We are very interested in the average time between spikes and its relation with matrix B (especially in cases where B^S might not have an inverse). Other characteristics of great interest are stationary distributions of membrane potentials and rate of convergence to a stationary distribution. It is particular intriguing for us to explore the cases where the system becomes sensitive to distribution of signals $\xi_{ij}^{(k)}$ and values μ_i and b_{ij} do not completely define the stationary regime. We explore such possibilities in the case of two neurons with Poisson inputs $X_i(t)$ and signals $\xi_{ij}^{(k)}$ with a specifically chosen distribution. A distant case from queueing networks which illustrates sensitivity results can be found in Foss and Chernova (1998).

There is a variety of other directions to explore in the upcoming research. A very challenging direction is the analysis of functional limit theorems. We also may branch out from the questions of stationarity and explore large deviations. Additionally, we could explore optimisation problems (e.g., finding a matrix B which maximises certain function of the network).

SECTION A

Appendix I

§ A.1. Proof of Proposition 1.12

Let $\xi_1, \xi_2, \ldots$ be positive i.i.d. r.v.'s with a common c.d.f. F. Let $S_0 = 0$ and $S_k = \xi_1 + \ldots \xi_k$, $k \geq 1$. Let τ be a counting r.v. with c.d.f. G, independent of $\{\xi_k\}_{k=1}^{\infty}$. Denote the c.d.f. of S_τ as H. Let

$$\overline{F}(x) = 1 - F(x),\ x \in \mathbb{R}, \tag{284}$$

$$\widehat{F}(\lambda) = \mathbb{E}e^{-\lambda\xi_1} = \int_0^\infty e^{-\lambda x} dF(x),\ \lambda \geq 0. \tag{285}$$

Define $\overline{G}, \widehat{G}, \overline{H}$, and $\widehat{H}$ similarly. We assume that there exist slowly varying (at infinity) functions l_1 and l_2 such that

$$\overline{F}(x) \sim \frac{l_1(x)}{x^\alpha} \text{ and } \overline{G}(x) \sim \frac{l_2(x)}{x^\beta} \tag{286}$$

as $x \to \infty$, for $\alpha, \beta \in (0, 1)$. We show that $\overline{H}$ possesses a similar asymptotic behaviour. For that we use the following result.

Proposition A.1. ***(Corollary 8.1.7, Bingham, Goldie and Teugels (1987))*** *For $0 \leq \alpha \leq 1$, l slowly varying (at infinity), the following are equivalent:*

$$1 - \widehat{F}(\lambda) \sim \lambda^\alpha l\left(\frac{1}{\lambda}\right),\ \textit{as } \lambda \downarrow 0, \tag{287}$$

$$\begin{cases} \overline{F}(x) \sim \frac{l(x)}{x^\alpha\Gamma(1-\alpha)},\ \textit{as } x \to \infty, & \textit{if } 0 \leq \alpha < 1, \\ \int_0^x t dF(t) \sim \int_0^x \overline{F}(t)dt \sim l(x),\ \textit{as } x \to \infty, & \textit{if } \alpha = 1. \end{cases} \tag{288}$$

Using this result, we get

$$1 - \widehat{F}(\lambda) \sim \lambda^\alpha \Gamma(1-\alpha) l_1\left(\frac{1}{\lambda}\right) \text{ and } 1 - \widehat{G}(\lambda) \sim \lambda^\beta \Gamma(1-\beta) l_2\left(\frac{1}{\lambda}\right),\ \text{as } \lambda \downarrow 0. \tag{289}$$

Let us analyse $\widehat{H}$:

$$\widehat{H}(\lambda) = \mathbb{E}e^{-\lambda S_\tau} = \sum_{k=1}^{\infty} e^{-\lambda(\xi_1+\ldots+\xi_k)}\mathbb{P}\{\tau = k\} = \mathbb{E}\left(\mathbb{E}e^{-\lambda\xi_1}\right)^\tau = \widehat{G}(-\ln\widehat{F}(\lambda)). \tag{290}$$

Since

$$-\ln\widehat{F}(\lambda) = -\ln(1-(1-\widehat{F}(\lambda))) \sim 1-\widehat{F}(\lambda), \text{ as } \lambda\downarrow 0, \tag{291}$$

we have

$$\begin{aligned}1-\widehat{H}(\lambda) &\sim 1-\widehat{G}\left(\lambda^\alpha\Gamma(1-\alpha)l_1\left(\frac{1}{\lambda}\right)\right) \\ &\sim \lambda^{\alpha\beta}\Gamma^\beta(1-\alpha)\Gamma(1-\beta)l_1^\beta\left(\frac{1}{\lambda}\right)l_2\left(\frac{1}{\lambda^\alpha\Gamma(1-\alpha)l_1\left(\frac{1}{\lambda}\right)}\right),\end{aligned} \tag{292}$$

as $\lambda\downarrow 0$, and finally

$$\overline{H}(x) \sim x^{-\alpha\beta}\frac{\Gamma^\beta(1-\alpha)\Gamma(1-\beta)}{\Gamma(1-\alpha\beta)}l_1^\beta(x)\,l_2\left(\frac{x^\alpha}{\Gamma(1-\alpha)l_1(x)}\right), \text{ as } x\to\infty. \tag{293}$$

§ A.2. Limit theorems for CTRW's

Let $\{(Y_n, J_n)\}_{n=1}^\infty$ be i.i.d. with (Y, J) on $(\mathbb{R}, \mathbb{R}_+)$ and set

$$S(n) = \sum_{k=1}^n Y_k \text{ and } T(n) = \sum_{k=1}^n J_k. \tag{294}$$

Denote

$$N(t) = \max\{n\geq 0:\ T(n)\leq t\}. \tag{295}$$

Process $(N(t),\ S(N(t)))$ is a *compound renewal process* with renewal times $T(n)$ and marks Y_n.

For the first result we assume

$$\mathbb{E}Y_1 = 0,\ \mathbb{E}Y_1^2 = 1 \tag{296}$$

and

$$\mathbb{P}\{J_1 > x\} \sim 1/(\Gamma(1-\alpha)x^\alpha L(x)), \text{ as } x\to\infty, \tag{297}$$

where $\alpha\in(0,1)$ and function L is slowly varying at infinity. Now we can formulate the first result.

Proposition A.2. *(Theorem 5.1, Kasahara (1984)) Suppose (296) and (297) ar satisfied for* $0 < \alpha < 1$. *Then*

$$\left\{ (\lambda^\alpha L(\lambda))^{-1/2} S(T(\lambda t)),\ t \geq 0 \right\} \overset{\mathcal{D}}{\Rightarrow} \left\{ B(l_\alpha(t)),\ t \geq 0 \right\},\ \text{as } \lambda \to \infty, \tag{298}$$

where l_α *is the inverse process of the one-sided stable process with Laplace transform* e^{-ts^α} *and* $B(\cdot)$ *is a Brownian motion which is independent of* l_α.

For the next two results we assume that there exist functions $B(c) > 0$ and $b(c) > 0$, $c \geq 0$, and a r.v.s A and D having a stable distribution such that

$$(B(n)S(n), b(n)T(n)) \Rightarrow (A, D),\ \text{as } n \to \infty. \tag{299}$$

We may assume that $B(\cdot)$ and $b(\cdot)$ are regularly varying (at infinity) functions. Then there exists a regularly varying function $\widetilde{b}$ such that $1/b(\widetilde{b}(c)) \sim c$, as $c \to \infty$. Define $\widetilde{B}(c) = B(\widetilde{b}(c))$.

Let $\{(A(t), D(t))\}_{t \geq 0}$ denote a stochastic process with independent increments such that $(A(1), D(1))$ has the same distribution as (A, D) (or *Lévy process* generated by (A, D)). Let $E(s) = \inf\{t \geq 0 :\ D(t) > s\}$. Finally, we can formulate the remaining two auxiliary result.

Proposition A.3. *(Corollary 3.4, Meerschaert and Scheffler (2004)) Assume* $b(n)T(n) \Rightarrow D$, *as* $n \to \infty$. *As* $c \to \infty$,

$$\left\{ \frac{N(ct)}{\widetilde{b}(c)},\ t \geq 0 \right\} \overset{\mathcal{D}}{\Rightarrow} \left\{ E(t), t \geq 0 \right\}. \tag{300}$$

Proposition A.4. *(Theorem 3.1, Jurlewicz et al. (2010)) Assume (299) holds. Then*

$$\left\{ \widetilde{B}(c)S(N(ct) + 1),\ t \geq 0 \right\} \overset{\mathcal{D}}{\Rightarrow} \left\{ A(E(t)), t \geq 0 \right\},\ \text{as } c \to \infty \tag{301}$$

and

$$\left\{ \widetilde{B}(c)S(N(ct)),\ t \geq 0 \right\} \overset{\mathcal{D}}{\Rightarrow} \left\{ A(E(t)-), t \geq 0 \right\},\ \text{as } c \to \infty, \tag{302}$$

where the scaling limit in the last convergence has to be interpreted as the right-continuous version of $\{A(E(t)-), t \geq 0\}$.

SECTION B

Appendix I

§ B.1. Proof of Remark 5.7

We prove the remark by assuming that the system of equations $\mathbf{x}B = \mathbf{1}$ has a solution, and then we prove that the solution is unique and has all positive coordinates. Let us rewrite the system $\mathbf{x}B = \mathbf{1}$ as

$$\sum_{j=1}^{N} x_j b_{ji} = H_i x_i + \sum_{j\neq i} w_j x_j = (H_i - w_i)x_i + \sum_{j=1}^{N} w_j x_j = 1, \quad i = 1, \dots, N. \quad (303)$$

Denote $M = \sum_{j=1}^{N} w_j x_j$ and get

$$x_i = \frac{1-M}{H_i - w_i} \quad i = 1, \dots, N \;\Rightarrow\; M = (1-M)\sum_{i=1}^{N} \frac{w_i}{H_i - w_i}$$
$$\Rightarrow\; M = \frac{\sum_{i=1}^{N} \frac{w_i}{H_i - w_i}}{1 + \sum_{i=1}^{N} \frac{w_i}{H_i - w_i}} < 1. \quad (304)$$

Thus, M and, therefore, $\mathbf{x}$ are uniquely defined through $\{H_i, w_i\}_{i=1}^{n}$ and $x_i > 0$, $i = 1, \dots, N$. The rest of the proof follows from the proof of Lemma 5.8.

§ B.2. Proof of Lemma 6.4

First, take $X(t)$, $t \in [0, 1]$, a standard one-dimensional Brownian motion. Then for the Brownian bridge with a bridge $B_x(t) = xt + B_0(t)$ and $x \leq y$ we have

$$\begin{aligned}\mathbb{P}\{\inf_{0\leq t\leq 1}(X(t)) \geq -k \mid X(1) = x\} &= \mathbb{P}\{\inf_{0\leq t\leq 1}(xt + B_0(t)) \geq -k\} \\ &\leq \mathbb{P}\{\inf_{0\leq t\leq 1}(yt + B_0(t)) \geq -k\} \\ &= \mathbb{P}\{\inf_{0\leq t\leq 1}(X(t)) \geq -k \mid X(1) = y\}. \quad (305)\end{aligned}$$

For a general N-dimensional Brownian motion $X(t)$, $t \in [0, 1]$, with a non-singular covariance matrix Σ, there exists an invertible matrix L and a vector $\mathbf{v}$ such that

$W(t) = LX(t) + \mathbf{v}t$ is a vector of N independent standard Brownian motions. Let $B_{\mathbf{0}}(t)$ denote the corresponding N-dimensional Brownian bridge. Denote $A_k = [-k, \infty)^N$. Then for $\mathbf{x}, \mathbf{y}$, such that $x_i \leq y_i$, we have

$$\begin{aligned}
&\mathbb{P}\left\{\min_{1\leq i\leq N} \inf_{0\leq t\leq 1}(X_i(t)) \geq -k |\ X(1) = \mathbf{x}\right\} \\
&= \mathbb{P}\left\{X(t) \in A_k,\ t \in [0,1] |\ X(1) = \mathbf{x}\right\} \\
&= \mathbb{P}\left\{W(t) \in LA_k + \mathbf{v}t,\ t \in [0,1] |\ W(1) = L\mathbf{x} + \mathbf{v}\right\} \\
&= \mathbb{P}\left\{L\mathbf{x}t + \mathbf{v}t + B_{\mathbf{0}}(t) \in LA_k + \mathbf{v}t,\ t \in [0,1]\right\} \\
&= \mathbb{P}\left\{\mathbf{x}t + L^{-1}B_{\mathbf{0}}(t) \in [-k, \infty)^N,\ t \in [0,1]\right\} \\
&\leq \mathbb{P}\left\{\mathbf{y}t + L^{-1}B_{\mathbf{0}}(t) \in [-k, \infty)^N,\ t \in [0,1]\right\} \\
&= \mathbb{P}\left\{\min_{1\leq i\leq N} \inf_{0\leq t\leq 1}(X_i(t)) \geq -k |\ X(1) = \mathbf{y}\right\}.
\end{aligned} \tag{306}$$

Using the properties of the Brownian bridge, one can replace the $\mathbf{y}$ with a measurable set Δ and prove the statement.

References

[1] Abiteboul, S., Preda, M., Cobena, G. (2003). Adaptive on-line page importance computation. *WWW '03: Proceedings of the 12th international conference on World Wide Web*, 280–290.

[2] Aldous, D., Fill, J. (2002). *Reversible Markov chains and random walks on graphs*, Available at http://www.stat.berkeley.edu/users/aldous/RWG/book.html.

[3] Alsmeyer, G., Sgibnev, M. (1999). On the tail behaviour of the supremum of random walk defined on a Markov chain. *Yokohama Math. J.*, **46,** 139–159.

[4] Arndt, K. (1980). Asymptotic properties of the distribution of the supremum of a random walk on a Markov chain. *Theory Probab. Appl.*, **25,** 309–324.

[5] Asmussen, S. (1989). Aspects of matrix Wiener-Hopf factorisation in applied probability. *Math. Scientist*, **14,** 101–116.

[6] Asmussen, S. (1991). Ladder heights and the Markov-modulated M/G/1 queue. *Stoch. Proc. Appl.*, **37,** 313–326.

[7] Asmussen, S. (2003). *Applied probability and queues, 2nd edition*, New-York: Springer-Verlag.

[8] Asmussen, S., Foss, S. (2014). On exceedance times for some processes with dependent increments. *J. Appl. Probab.*, **51,** 136–151.

[9] Asmussen, S., Henriksen, L. F., Kloppelberg, C. (1994). Large claims approximations for risk processes in a Markovian environment. *Stochactic Process. Appl.*, **54,** 29–43.

[10] Asmussen, S., Turova, T. S. (1998). Stationarity properties of neural networks. *J. Appl. Probab.*, **35,** 783–794.

[11] Baeza-Yates, R. A., Culberson, J. C., Rawlins., G. J. E. (1993). Searching in the plane. *Information and Computation*, **106,** 234–252.

[12] Becker-Kern, P., Meerschaert, M. M. and Scheffler, H.-P. (2004). Limit theorems for coupled continuous time random walks. *Annals of Probability*, **32,** 730–756.

[13] Billingsley, P. (1968). *Convergence of probability measures.* John Wiley & Sons.

[14] Bingham, N. H., Goldie, C. M., Teugels, J. L. (1987). *Regular variation*, Encyclopedia of Mathematics and its Applications, **27**, Cambridge University Press.

[15] Borodin, A., Linial, N., Saks, M. (1992). An optimal online algorithm for metrical tasks systems. *Journal of ACM*, **39,** 745–763.

[16] Borovkov, A. A. (1965). On the first passage time for one class of processes with independent increments. *Theory Probab. Appl.*, **10**, 331–334.

[17] Borovkov, A. A. and Borovkov, K. A. (2008). *Asymptotic analysis of random walks: Heavy-tailed distributions*, Cambridge Univ. Press, Encyclopedia of Mathematics and Its Applications 118.

[18] Borovkov, A. A., Foss, S. G. (1992). Stochastically recursive sequences and their generalisations. *Siberian Adv. Math.*, **1**, 16–81.

[19] Borst, S., Jonckheere, M., Leskelä, L. (2008). Stability of parallel queueing systems with coupled service rates. *Discrete Event Dynamic Systems*, **18,** 447–472.

[20] Burkitt, A. (2006a). A review of the integrate-and-fire neuron model: 1. Homogeneous synaptic input. *Biol. Cybern.*, **95**, 1–19.

[21] Burkitt, A. (2006b). A review of the integrate-and-fire neuron model: 2. Inhomogeneous synaptic input and network properties. *Biol. Cybern.*, **95**, 97–112.

[22] Cáceres, M. J., Carrillo, J. A., Perthame, B. (2011). Analysis of nonlinear noisy integrate and fire neuron models: blow-up and steady states. *J. Math. Neurosci.*, **1**, 1–33.

[23] Cline, D. B. H. (1986) Convolution tails, product tails and domains of attraction. *Probab. Theory Related Fields*, **72,** 529–557.

[24] Coppersmith, D., Doyle, P., Raghavan, P., Snir, M. (1993). Random walks on weighted graphs and applications to on-line algorithms. *J. Assoc. Comput. Mach.*, **40,** 421–453.

[25] Cottrell, M. (1992). Mathematical analysis of a neural network with inhibitory coupling. *Stoch. Proc. Appl.*, **40,** 103–126.

[26] Dai, J. G. (1995). On positive Harris recurrence of multiclass queueing networks: a unified approach via fluid limit models . *Ann. Appl. Probab.*, **5**, 49–77.

[27] De Masi, A., Galves, A., Löcherbach, E., Presutti, E. (2015). Hydrodynamic limit for interacting neurons. *J. Stat. Phys.*, **158**, 866–902.

[28] Delarue, F., Inglis, J., Rubenthaler, S., Tanré, E. (2015). Global solvability of a networked integrate-and-fire model of McKean-Vlasov type. *Ann. Appl. Probab.*, **25**, 2096–2133.

[29] Denisov, D., Foss, S. and Korshunov, D. (2010). Asymptotics of randomly stopped sums in the presence of heavy tails. *Bernoulli*, **16,** 971–994.

[30] Dobrushin, R. L. (1955). Lemma on the limit of compound random functions. *Uspekhi Mat. Nauk*, **10,** 157–159.

[31] Embrechts, P., Goldie, C. M. (1982) On convolution tails. *Stochastic Process. Appl.*, **13,** 263–278.

[32] Feller, W. (1971a). *An Introduction to Probability Theory and Its Applications*, 1. John Wiley & Sons.

[33] Feller, W. (1971b). *An Introduction to Probability Theory and Its Applications*, 2. John Wiley & Sons.

[34] Foss, S. G., Konstantopoulos, T. (2004). An overview of some stochastic stability methods. *J. Oper. Res. Soc. JPN*, **47**, 275–303.

[35] Foss, S., Konstantopoulos, T., Zachary, S. (2007). Discrete and continuous time modulated random walks with heavy-tailed increments. *J. Theoret. Probab.*, **20,** 581–612.

[36] Foss, S., Shneer, S., Thomas, J. P., Worrall T. (2018). Stochastic stability of monotone economies in regenerative environments. *J. Econom. Theory*, **173,** 334–360.

[37] Foss, S., Shneer, S., Tyurlikov, A. (2012). Stability of a Markov-modulated Markov chain, with application to a wireless network governed by two protocols. *Stochastic Systems*, **2,** 208–231.

[38] Foss, S. and Zachary, S. (2003). The maximum on a random time interval of a random walk with long-tailed increments and negative drift. *Ann. Appl. Probab.*, **13,** 37–53.

[39] Fricker, C., Robert, P., Saada, E., Tibi, D. (1994). Analysis of some networks with interaction. *Ann. Appl. Probab.*, **4**, 1112–1128.

[40] Gamarnik, D. (2004). Stochastic bandwidth packing process: stability conditions via Lyapunov function technique. *Queueing systems*, **48,** 339–363.

[41] Gamarnik, D., Squillante, M. (2005). Analysis of stochastic online bin packing processes. *Stochastic Models*, **21,** 401–425.

[42] Georgiou, N., Wade, A. R. (2014). Non-homogeneous random walks on a semi-infinite strip. *Stochactic Process. Appl.*, **124,** 3179–3205.

[43] Gerstner, W., Kistler, W. (2002). *Spiking neuron models: Single neurons, populations, plasticity*, Cambridge: Cambridge University Press.

[44] Hansen, N. R., Jensen, A. T. (2005). The extremal behaviour over regenerative cycles for Markov additive processes with heavy tails. *Stochactic Process. Appl.*, **115,** 579–591.

[45] Hodgkin, A. L., Huxley, A. F. (1952). Quantitative description of ion currents and its applications to conduction and excitation in nerve membranes. *J. Physiol. (Lond.)*, **117**, 500–544.

[46] Inglis, J., Talay, D. (2015). Mean-field limit of a stochastic particle system smoothly interacting threshold hitting-times and applications to neural networks with dendritic component. *SIAM J. Math. Anal.*, **47**, 3884–3916.

[47] Jelenkovic, P. R., Lazar, A. A. (1998). Subexponential Asymptotics of a Markov-Modulated Random Walk with Queueing Applications. *J. Appl. Probab.*, **35,** 325–347.

[48] Jurlewicz, A., Kern, P., Meerschaert, M. M., Scheffler, H.-P. (2010). Oracle continuous time random walks. Available at http://www.uni-siegen.de/fb6/src/scheffler/research/octrw2.pdf.

[49] Karpelevich, F. I., Malyshev, V. A., Rybko, A. N. (1995). Stochastic evolution of neural networks. *Markov Process. Related Fields*, **1,** 141–161.

[50] Kasahara, Y. (1984). Limit theorems for Lévy processes and Poisson point processes and their applications to Brownian excursions. *J. Math. Kyoto University*, **24,** 521–538.

[51] Lapicque, L. (1907). Recherches quatitatives sur l'excitation electrique des nerfs traitée comme une polarization. *J. Physiol. (Paris)*, **9**, 620–635.

[52] Latham, P. E., Richmond, B. J., Nelson, P. G., Nirenberg, S. (2000). Intrinsic dynamics in neuronal networks. I. Theory. *J. Neurophysiol.*, **83**, 808–827.

[53] Leslie, J. (1989). On the non-closure under convolution of the class of subexponential distributions. *J. Appl. Probab.*, **26,** 58–66.

[54] Litvak, N., Robert, P. (2008). Analysis of an on-line algorithm for solving large Markov chains. *The 3rd International Workshop on Tools for solving Structured Markov Chains (Athens), ACM.*

[55] Litvak, N., Robert, P. (2012). A scaling analysis of a cat and mouse Markov chain. *Annals of Probability*, **22,** 792–826.

[56] Lu, Y., Li, S. (2005). On the probability of ruin in a Markov-modulated risk model. *Insurance: Mathematics and Economics*, **37,** 522–532.

[57] Malyshev, V. A. (1972). Classification of two-dimensional positive random walks and almost linear semimartingales. *Dokl. Akad. Nauk SSSR*, **202,** 526–528 *(in Russian)*.

[58] Malyshev, V. A. (1993). Networks and dynamical systems. *Adv. Appl. Prob.*, **26,** 140–175.

[59] Malyshev, V. A., Ignatyuk, I. A. (1993). Classification of random walks in Z_+^4. *Adv. Appl. Prob.*, **25,** 140–175.

[60] Malyshev, V. A., Menshikov, M. V. (1979). Ergodicity, continuity and analyticity of countable Markov chains. *Tr. Mosk. Mat. Obs.*, **39,** 3–48 *(in Russian)*.

[61] Manasse, M. S., McGeoch, L. A., Sleator, D. D (1990). Competitive algorithms for server problems. *Journal of Algorithms*, **11,** 208–230.

[62] Meerschaert, M. M. and Scheffler, H.-P. (2001). *Limit distributions for sums of independent random vectors*. John Wiley, New York.

[63] Meerschaert, M. M. and Scheffler, H.-P. (2004). Limit theorems for continuous-time random walks with infinite mean waiting times. *J. Appl. Probab.*, **41,** 623–638.

[64] Prasolov, T. (2018). Stochastic stability of a system of perfect integrate-and-fire inhibitory neurons. *(Submitted for review)*.

[65] Prasolov, T., Foss, S., Shneer, S. (2018). On the Cat-and-Mouse Markov chain and its generalisations. *(Submitted for review)*.

[66] Robert, P., Touboul, J. (2016). On the dynamics of random neuronal networks. *J. Stat. Phys.*, **165**, 545–584.

[67] Rybko, A. N., Stolyar, A. L. (1992). Ergodicity of stochastic processes describing the operations of open queueing networks. *Probl. Peredachi Inf.*, **28**, 3–26.

[68] Sato, K. (1999). *Lévy processes and infinitely divisible distributions*, Cambridge studies in advanced mathematics, **68**.

[69] SHAH, D., SHIN, J. (2012). Randomized scheduling algorithm for queueing networks. *Ann. Appl. Probab.*, **22,** 128–171.

[70] SKOROKHOD, A. V. (1956). Limit theorems for stochastic processes. *Theory Probab. Appl.*, **1,** 261–290.

[71] SPITZER, F. (1964). *Principles of random walk*, Springer, New York.

[72] STEIN, R. B. (1965). A theoretical analysis of neuronal variability. *Biophys. J.*, **5**, 173–194.

[73] STOLYAR, A. I. (1995). On the stability of multiclass queueing networks: A relaxed sufficient condition via limiting fluid processes. *Markov Process. Relat.*, **1**, 491–512.

[74] TAILLEFUMIER, T.. Informal talk.

[75] TAILLEFUMIER, T., TOUBOUL, J., MAGNASCO, M. (2012). Exact event-driven implementation for recurrent networks of stochastic perfect integrate-and-fire neurons. *Neural Comput.*, **24**, 3145–3180.

[76] UCHIYAMA, K. (2011a). The first hitting time of a single point for random walks. *Electron. J. Probab.*, **16,** 1960–2000.

[77] UCHIYAMA, K. (2011b). One dimensional lattice random walks with absorption at a point/on a half line. *J. Math. Soc. Japan*, **63,** 675–713.

[78] WANG, B. C., LIU, Y. Y. (2011). Local asymptotics of a Markov modulated random walk with heavy-tailed increments. *Acta Mathematica Sinica, English Series*, **27,** 1843–1854.